

UNIVERSIDADE FEDERAL DE ALAGOAS  
INSTITUTO DE COMPUTAÇÃO  
PROGRAMA DE PÓS GRADUAÇÃO EM MODELAGEM COMPUTACIONAL DE  
CONHECIMENTO

DAVID MEDEIROS BATINGA

**Um Modelo Computacional para a Classificação Automática do Tipo de Personalidade  
de Estudantes do Ensino a Distância**

Maceió-AL  
Janeiro de 2020



DAVID MEDEIROS BATINGA

**Um Modelo Computacional para a Classificação Automática do Tipo de Personalidade de Estudantes do Ensino a Distância**

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre pelo Programa de Pós-Graduação em Modelagem Computacional de Conhecimento do Instituto de Computação da Universidade Federal de Alagoas.

Orientador: Dr. Ranilson Oscar Araújo Paiva  
Coorientador: Dr. Jorge Artur Peçanha de Miranda Coelho

Maceió-AL

Janeiro de 2020



**Catálogo na fonte**  
**Universidade Federal de Alagoas**  
**Biblioteca Central**  
**Divisão de Tratamento Técnico**

Bibliotecária: Taciana Sousa dos Santos – CRB-4 – 2062

B333m Batinga, David Medeiros.

Um modelo computacional para a classificação automática do tipo de personalidade de estudantes do ensino à distância / David Medeiros  
Batinga. – 2020.

83 f. : il. figs.; tabs. color.

Orientador: Ranilson Oscar Araújo Paiva.

Coorientador: Jorge Artur Peçanha de Miranda Coelho.

Dissertação (Mestrado em Modelagem Computacional do Conhecimento) – Universidade Federal de Alagoas. Instituto de Computação. Maceió, 2020.

Bibliografia: 67-72.

Anexos: f. 73-83.

1. Traços de personalidade. 2. Aprendizagem. 3. Educação à distância.  
4. Modelo computacional. I. Título.

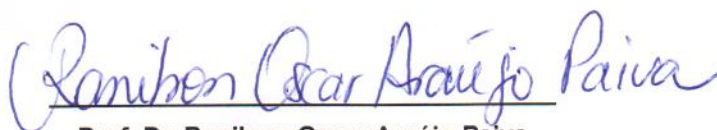
CDU: 004.8: 159.923.3

**Folha de Aprovação**

David Medeiros Batinga

Um Modelo Computacional para a Classificação Automática do Tipo de Personalidade  
de Estudantes do Ensino a Distância

Dissertação submetida ao corpo docente do  
Programa de Pós-Graduação em Modelagem  
Computacional de Conhecimento da  
Universidade Federal de Alagoas e aprovada  
em 15 de janeiro de 2020.



**Prof. Dr. Ranilson Oscar Araújo Paiva**

Instituto de Computação – UFAL

Orientador



**Prof. Dr. Jorge Artur Peçanha de Miranda Coelho**

Faculdade de Medicina – UFAL

Coorientador

**Banca Examinadora:**



**Prof. Dr. Diego Dermeval Medeiros da Cunha Matos**

Faculdade de Medicina – UFAL

Examinador Interno



**Prof. Dr. Alan Pedro da Silva**

Instituto de Computação – UFAL

Examinador externo

## RESUMO

Nas últimas décadas, o papel dos traços de personalidade tem ganho relevância para compreender, explicar, e mais recentemente, prever padrões comportamentais que os indivíduos externalizam quando vivenciam diversas situações cotidianas, incluindo: tomar decisões, estabelecer relações interpessoais, lidar com questões financeiras, gerir carreira e trabalho, etc. Pesquisas têm demonstrado a importância e a aplicabilidade do conceito de personalidade no contexto educacional; tendo sido identificadas correlações entre a personalidade dos estudantes com vários assuntos de interesse para o processo de aprendizagem. Por exemplo, estudos anteriores identificaram uma forte correlação entre a personalidade dos estudantes e as pontuações obtidas por esses alunos nas provas de avaliação de aprendizagem; sendo possível, ao conhecer antecipadamente o tipo de personalidade de um aluno, estimar com uma boa margem de segurança se ele obterá aprovação nas avaliações do seu curso. Entretanto, o diagnóstico dos traços de personalidade de um aluno não é uma atividade simples e geralmente é realizado por meio de entrevistas e observações, ou através do uso de questionários. Num ambiente de ensino à distância, a aplicação de ambas abordagens se mostra ainda mais complexa, então pesquisas têm buscado novos caminhos para viabilizar o diagnóstico dos traços de personalidade. Uma das alternativas é realizar o diagnóstico da personalidade com o auxílio de modelos matemáticos que são aplicados para análise dos comportamentos e dos padrões de interação dos alunos realizados no ambiente de ensino à distância. O objetivo deste trabalho foi investigar a possibilidade de desenvolvimento e avaliação de um modelo computacional que se propõe a identificar os traços de personalidade de alunos através da análise das interações realizadas no ambiente virtual de aprendizagem. Para isso, delineou-se uma pesquisa experimental realizada no âmbito dos cursos de graduação à distância da Universidade Federal de Alagoas. No experimento, um questionário de personalidade foi aplicado a uma amostra de estudantes e, na sequência, um conjunto de algoritmos de classificação foi usado para analisar o uso do ambiente de aprendizagem, a fim de se obter possíveis correlações entre os traços de personalidade dos alunos e as suas interações no ambiente virtual de aprendizagem. Como resultado geral, foi observado que nenhum dos algoritmos de classificação obteve resultados satisfatórios.

**Palavras-chaves:** Educação a Distância; Traços de Personalidade; Classificação da Personalidade; Mineração de Dados Educacionais.





## ABSTRACT

In the last decades, the role of personality traits has gained relevance for understanding, explaining and more recently, the behavioral patterns that people exhibit when they experience different situations, including everyday situations such as: making decisions, using interpersonal relationships, dealing with finance, career and work management issues, etc. Research has demonstrated the importance and applicability of the concept of personality in the educational context; correlations were identified between the students' personality with various subjects of interest to the learning process. For example, previous studies have identified a strong correlation between students' personality and the scores obtained by those students in the learning assessment tests; it is possible, when knowing in advance the personality type of a student, to estimate with a good safety margin whether he will pass his course evaluations. However, the diagnosis of a student's personality traits is not a simple activity and is usually carried out through interviews and observations, or through the use of questionnaires. In a distance learning environment, the application of both approaches proves to be even more complex, so research has sought new ways to make the diagnosis of personality traits viable. One of the alternatives is to carry out the diagnosis of the personality with the aid of mathematical models that are applied to analyze the behaviors and interaction patterns of students carried out in the distance learning environment. The objective of this work was to investigate the possibility of developing and evaluating a computational model that proposes to identify the personality traits of students through the analysis of the interactions carried out in the virtual learning environment. For this, an experimental research was carried out in the context of distance undergraduate courses at the Federal University of Alagoas. In the experiment, a personality questionnaire was applied to a sample of students and, subsequently, a set of classification algorithms was used to analyze the use of the learning environment, in order to obtain possible correlations between students' personality traits. and their interactions in the virtual learning environment. As a general result, it was observed that none of the classification algorithms obtained satisfactory results.

**Keywords:** Distance Education; Personality traits; Personality Classification; Educational Data Mining.



## LISTA DE TABELAS

Tabela 1 – As dimensões (preferências) e as definições do modelo de personalidade classificação tipológica de Myers-Briggs . . . . .	27
Tabela 2 – Atributos dos alunos definidos pela pesquisa de Halawa, Shehab e Hamed (2015) e recuperados a partir do ambiente de aprendizagem da Universidade Alemã do Cairo . . . . .	28
Tabela 3 – Atributos dos alunos relacionados às interações da rede social escolhida pela pesquisa de Halawa, Shehab e Hamed (2015) . . . . .	28
Tabela 4 – Atributos dos alunos recuperados por meio das interações com o ambiente de ensino edX durante as atividades do curso <i>Data Analysis</i> e utilizados na pesquisa de Chen et al. (2016) . . . . .	30
Tabela 5 – Atributos dos alunos relacionados às interações no ambiente virtual de aprendizagem escolhidos pela pesquisa de Ghorbani e Montazer (2015) . . . . .	32
Tabela 6 – Tabela comparativa dos trabalhos relacionados . . . . .	33
Tabela 7 – Distribuição de alunos participantes do experimento por curso e gênero . . . . .	37
Tabela 8 – Caracterização descritiva das pontuações dos traços de personalidade dos participantes do experimento . . . . .	39
Tabela 9 – Atributos utilizados pela pesquisa e que são recuperados a partir do sistema de gestão de aprendizagem da UFAL . . . . .	40
Tabela 10 – Atributos utilizados pela pesquisa e que são obtidos a partir das respostas do questionário aplicado . . . . .	41
Tabela 11 – Experimento - Definição dos níveis de fatores . . . . .	43
Tabela 12 – Experimento - Definição formal das hipóteses . . . . .	44
Tabela 13 – Definição das execuções dos ensaios do experimento . . . . .	45
Tabela 14 – Resultados das métricas de precisão para os algoritmos avaliados pelo experimento para o traço de personalidade Abertura . . . . .	58
Tabela 15 – Testes de Hipóteses dos resultados dos algoritmos para predição da pontuação de Traço de Personalidade Abertura . . . . .	58
Tabela 16 – Resultados das métricas de precisão para os algoritmos avaliados pelo experimento para o traço de personalidade Conscienciosidade . . . . .	59
Tabela 17 – Testes de Hipóteses dos resultados dos algoritmos para predição da pontuação de Traço de Personalidade Conscienciosidade . . . . .	59
Tabela 18 – Resultados das métricas de precisão para os algoritmos avaliados pelo experimento para o traço de personalidade Extroversão . . . . .	60
Tabela 19 – Testes de Hipóteses dos resultados dos algoritmos para predição da pontuação de Traço de Personalidade Extroversão . . . . .	60
Tabela 20 – Resultados das métricas de precisão para os algoritmos avaliados pelo experimento para o traço de personalidade Amabilidade . . . . .	61

Tabela 21 – Testes de Hipóteses dos resultados dos algoritmos para predição da pontuação de Traço de Personalidade Amabilidade . . . . .	61
Tabela 22 – Resultados das métricas de precisão para os algoritmos avaliados pelo experimento para o traço de personalidade Neuroticismo . . . . .	62
Tabela 23 – Testes de Hipóteses dos resultados dos algoritmos para predição da pontuação de Traço de Personalidade Neuroticismo . . . . .	63

## SUMÁRIO

1	<b>INTRODUÇÃO</b>	13
2	<b>FUNDAMENTAÇÃO TEÓRICA</b>	17
2.1	<b>Personalidade</b>	17
2.1.1	<b>Teorias de Personalidade</b>	17
2.1.2	<b>Teoria dos Traços da Personalidade</b>	19
2.1.3	<b>Origem e estabilidade da personalidade</b>	20
2.1.4	<b>Modelos de Traços da Personalidade</b>	20
2.1.5	<b>Modelo dos Cinco Grandes Fatores da Personalidade</b>	21
2.1.5.1	<b>Inventário dos Cinco Grandes Fatores da Personalidade (IGFP-5) – <i>The Big Five Inventory</i></b>	23
2.1.6	<b>O impacto das diferenças individuais na aprendizagem</b>	24
2.2	<b>Mineração de Dados Educacionais</b>	24
3	<b>TRABALHOS RELACIONADOS</b>	27
3.1	<b>Predicting Student Personality Based on a Data-Driven Model from Student Behavior on LMS and Social Networks</b>	27
3.2	<b>On the Impact of Personality in Massive Open Online Learning</b>	29
3.3	<b>E-learners' personality identifying using their network behaviors</b>	31
3.4	<b>Tabela Comparativa</b>	32
4	<b>PROPOSTA DA PESQUISA</b>	35
4.1	<b>Construção da base de dados</b>	36
4.1.1	<b>Seleção da Amostra</b>	36
4.1.2	<b>Escolha do Instrumento de Avaliação de Personalidade</b>	37
4.1.3	<b>Coleta de Registros (<i>Logs</i>) das Interações no Ambiente Virtual de Aprendizagem</b>	38
4.1.4	<b>Construção do conjunto de dados do Experimento</b>	39
4.2	<b>Avaliação de Algoritmos Regressão para a Classificação de Traços de Personalidade</b>	41
4.2.1	<b>Questões de investigação e Hipóteses</b>	41
4.2.2	<b>Fatores e Variáveis de Respostas</b>	43
4.2.3	<b>Níveis dos fatores</b>	43
4.2.4	<b>Definição formal das hipóteses</b>	43
4.2.5	<b>Design do Experimento</b>	44
4.2.6	<b>Medidas de precisão</b>	44
4.2.7	<b>Execução dos Algoritmos</b>	48

4.2.7.1	Regressão Linear . . . . .	49
4.2.7.2	<i>Instance-Based Learning</i> (IBL) . . . . .	50
4.2.7.3	M5Rules . . . . .	50
4.2.7.4	M5P . . . . .	51
4.2.7.5	Rede Neural Perceptron Multicamadas . . . . .	52
4.2.7.6	<i>Support Vector Machines</i> (SVM) . . . . .	54
5	<b>RESULTADOS</b> . . . . .	57
5.1	Traço de Personalidade Abertura . . . . .	57
5.2	Traço de Personalidade Conscienciosidade . . . . .	59
5.3	Traço de Personalidade Extroversão . . . . .	60
5.4	Traço de Personalidade Amabilidade . . . . .	61
5.5	Traço de Personalidade Neuroticismo . . . . .	62
6	<b>CONCLUSÃO</b> . . . . .	65
6.1	Limitações . . . . .	66
6.2	Trabalhos Futuros . . . . .	66
	<b>REFERÊNCIAS</b> . . . . .	69
	<b>ANEXO A – QUESTIONÁRIO INVENTÁRIO DOS CINCO GRAN-</b> <b>DES FATORES DE PERSONALIDADE</b> . . . . .	75
	<b>ANEXO B – AUTORIZAÇÃO INSTITUCIONAL PARA REALIZA-</b> <b>ÇÃO DA PESQUISA – CIED</b> . . . . .	77
	<b>ANEXO C – PARECER CONSUBSTANCIADO DO CEP</b> . . . . .	79

## 1 INTRODUÇÃO

Pesquisas indicam que os traços de personalidade influenciam vários aspectos da vida cotidiana e exercem um importante papel para descrever padrões de comportamento que uma pessoa demonstra, quando confrontada com diversos estímulos e situações; incluindo, entre outros, sua percepção acerca da dedicação e realização nos estudos e no trabalho (JUDGE et al., 1999); sua facilidade em estabelecer relacionamentos interpessoais, resiliência a dificuldades e desafios (OSHIO et al., 2018), e concepção sobre a sua autoestima e felicidade (OZER; BENET-MARTÍNEZ, 2006).

O estudo da personalidade passou por um processo de estruturação e formalização na área da Psicologia a partir da década de 1930. Um dos marcos desse processo foi a publicação da obra de Allport (1937), que definiu traço de personalidade como um sistema neuropsíquico, intrínseco a cada indivíduo, capaz de receber e converter muitos estímulos funcionalmente equivalentes em padrões de comportamento congruentes. Em outras palavras, os traços de personalidade são predisposições a responder igualmente ou de modo semelhante a determinados estímulos.

A presente pesquisa tem como foco investigar o conceito de traços de personalidade no âmbito da educação, mais precisamente na modalidade de ensino à distância (EAD). Trabalhos anteriores demonstraram que o conceito de personalidade é de relevância prática para a área educacional, tendo sido provadas correlações entre os traços de personalidade com temas como: modelo e agrupamento de alunos, desempenho acadêmico (FEYTER et al., 2012; PAUNONEN; ASHTON, 2013), nível de engajamento e dedicação dos alunos (KOMARRAJU; KARAU; SCHMECK, 2009).

A análise de traços de personalidade de indivíduos é realizada, tradicionalmente, por meio da aplicação de questionários psicométricos (ANDRADE, 2008). Em se tratando da modalidade EAD, a utilização dessa abordagem se torna complexa e dispendiosa, tendo em vista que não existindo restrição de espaço físico, os cursos a distância podem conter um número elevado de alunos (ASSUMPÇÃO, 2017). Para contornar essa dificuldade, há a possibilidade de se adotar o envio de questionários por meio eletrônico, mas, em geral, o percentual de retorno para esse tipo de contato é baixo (MANFREDA et al., 2008). Um outro problema associado ao uso de questionários é que há margem para o preenchimento errado (deliberado ou não) dos questionários; possibilitando respostas influenciadas pelo interesse em ampliar ou em omitir características do respondente (GROVES, 1989).

Esta pesquisa avaliar a possibilidade de implementação de um modelo computacional, apoiado pela teoria de traços de personalidade, que permita a classificação automática do tipo de personalidade dos alunos por meio de técnicas de mineração de dados educacionais.

Segundo a proposta, essa análise será baseada exclusivamente nos registros dos hábitos, ações e interações realizados pelos alunos durante as atividades pedagógicas realizadas num

Ambiente Virtual de Aprendizagem (AVA), mais especificamente em Sistemas de Gestão de Aprendizagem<sup>1</sup>. Algoritmos de aprendizagem de máquina podem ser usados no decorrer das atividades pedagógicas dos cursos para analisar as interações dos alunos, com o objetivo de estabelecer correlações entre o uso da plataforma de ensino e seus traços de personalidade. Seria, então, possível identificar o tipo de personalidade dos alunos de forma automática, tornando dispensável a aplicação de questionários psicométricos.

O diagnóstico dos tipos de personalidade dos estudantes pode auxiliar docentes e tutores nas tomadas de decisão a cerca de intervenções pedagógicas que tornem a aprendizagem dos alunos mais eficiente.

A identificação dos traços de personalidade dos alunos também abre a oportunidade de implementar funcionalidades e mecanismos no Ambiente Virtual de Aprendizagem que viabilizem a personalização e individualização do ensino, de modo que o processo de instrução seja adaptado para cada estudante, em função do seu estilo de personalidade.

Nossa hipótese é que, através dos dados educacionais referentes às interações dos estudantes no LMS, sejamos capazes de detectar padrões (comportamentos recorrentes) que nos permitirão determinar os traços de personalidade dos estudantes.

O objetivo é que este modelo sirva como ferramenta prática para que professores, tutores e equipe pedagógica possam identificar os traços de personalidade dos alunos, viabilizando, assim, a implementação de melhorias nas intervenções pedagógicas personalizadas, que objetivem o incremento na aprendizagem.

A pesquisa vai transcorrer de acordo com a seguinte metodologia: num primeiro momento, um questionário de diagnóstico de personalidade será aplicado a uma amostra de alunos matriculados em cursos da modalidade de ensino a distância. Na etapa seguinte, alguns parâmetros relacionados às interações dos alunos serão extraídos do LMS. Finalmente, os parâmetros das interações serão analisados e comparados com os resultados do questionário aplicado, a fim de se buscar identificar correlações entre os traços de personalidade dos alunos e os padrões de uso do ambiente de aprendizagem. Diferentes algoritmos de aprendizagem de máquina serão usados para tentar encontrar possíveis correlações.

Ao final do experimento, busca-se comprovar ou refutar a existência de uma correlação significativa entre os traços de personalidade dos alunos e o padrão de hábitos e interações que estes demonstram ao utilizar o ambiente de aprendizagem.

Esta proposta de dissertação está estruturada em 5 capítulos. No capítulo 2 é apresentada uma fundamentação teórica sobre os temas e conhecimentos abordados e que serão úteis para o desenvolvimento desta dissertação. No capítulo 3 são discutidos trabalhos relacionados à temática de classificação de personalidade de estudantes no contexto de ensino à distância. O

---

<sup>1</sup> do inglês: Learning Management System (LMS). Os sistemas Moodle ([www.moodle.org](http://www.moodle.org)), Blackboard ([www.blackboard.com/learning-management-system](http://www.blackboard.com/learning-management-system)) e Canvas ([www.canvaslms.com](http://www.canvaslms.com)) são alguns exemplos de sistemas LMS.



capítulo 4 expõe a proposta da pesquisa e sua contribuição. E por fim, o capítulo 5 descreve a metodologia que será seguida durante o processo de construção do modelo computacional a ser proposto por este trabalho.



## **2 FUNDAMENTAÇÃO TEÓRICA**

Nesse capítulo, será apresentada a fundamentação teórica que serviu de base para esta pesquisa, onde serão abordados os conceitos necessários para o entendimento desta dissertação. Nesse sentido, as seções estão organizadas da forma como segue: a seção 2.1 fornecerá uma visão geral sobre o conceito de personalidade e a seção 2.2 que realiza uma introdução sobre mineração de dados educacionais.

### **2.1 Personalidade**

As pessoas diferem umas das outras, tanto física como psicologicamente. Certos indivíduos tendem a ser mais extrovertidos e anseiam por estabelecer interações sociais, enquanto outros são introvertidos e mais reservados; algumas pessoas são calmas e equilibradas, em contrapartida outras mostram-se frequentemente ansiosas. Essas diferenças individuais, relacionadas a tendências comportamentais, estão associadas ao conceito de personalidade (FEIST; FEIST; TOMI-ANN, 2015).

Pesquisas indicam que os traços de personalidade influenciam vários aspectos da vida cotidiana e exercem um importante papel para descrever padrões de comportamento que uma pessoa demonstra, quando confrontada com diversos estímulos e situações; incluindo, entre outros, sua percepção acerca da dedicação e realização nos estudos e no trabalho (BARRICK; MOUNT, 1991; JUDGE et al., 1999; ROTHMANN; COETZER, 2003); sua facilidade em estabelecer relacionamentos interpessoais, resiliência a dificuldades e desafios (OSHIO et al., 2018), concepção sobre a sua autoestima e felicidade (OZER; BENET-MARTÍNEZ, 2006), ponto de vista sobre saúde e longevidade (BOGG; ROBERTS, 2004; BOGG; ROBERTS, 2012; JACKSON et al., 2015) e julgamento a cerca de decisões financeiras (LAURIOLA; LEVIN, 2001; BIBBY; FERGUSON, 2011).

#### **2.1.1 Teorias de Personalidade**

No cotidiano o termo personalidade apresenta vários significados, sendo empregado em diferentes contextos, mas a maioria desses significados se enquadra em um ou dois cenários de uso. O primeiro uso vincula o termo à ideia de atração social ou de habilidades sociais. Nesse cenário, a personalidade de um indivíduo está associada a sua capacidade de produzir reações positivas ou negativas nas outras pessoas. Por exemplo, um aluno que constantemente se mostra disciplinado e dedicado em sala de aula, acaba por produzir uma reação positiva no seu professor. O segundo uso considera a personalidade de um indivíduo como a impressão mais marcante que este transmite aos demais. Nesse contexto, pode-se afirmar que uma pessoa tem uma “personalidade extrovertida”, “personalidade agressiva” ou “personalidade tímida”. Fica evidente a existência de um elemento de avaliação nas duas situações de uso, onde o juízo ocorre

naturalmente: o avaliador escolhe atributos e qualidades que ele considera particularmente mais evidentes no indivíduo e os adota para definir um arcabouço das reações e impressões geradas nas outras pessoas; e a partir daí o sujeito observador define o tipo de personalidade daquela pessoa (HALL; LINDZEY; CAMPBELL, 2007).

A limitação do sentido popular é que há obrigatoriedade de um avaliador para determinar a personalidade de um indivíduo. Segundo esse significado, a personalidade de uma pessoa é fruto unicamente da avaliação dos demais; sendo uma abordagem muito suscetível à subjetividade. Existe, então, a necessidade de buscar uma alternativa de conceituação que não seja exclusivamente dependente da avaliação de terceiros e que também se baseie nos atributos e qualidades do sujeito. Entretanto, não se estabeleceu, entre os teóricos, uma definição universal sobre personalidade e sua concepção varia em função da linha teórica de referência na qual esse constructo será estudado, como afirma Pasquali (2003) "[...] a expressão *personalidade* é tão ampla em seu significado ou, melhor, tão vaga, que praticamente cada psicólogo a entende do seu modo".

Allport (1937) identificou diversas definições de personalidade, dentre as quais, algumas merecem destaque. Uma das opções de definição de personalidade, chamada de definição biossocial, equipara-se com o uso popular do termo, descrito anteriormente, no qual a personalidade é classificada pela impressão social que o indivíduo provoca nas outras pessoas. Allport entra em desacordo com a definição biossocial e afirma ser preferível a chamada definição biofísica, onde a personalidade deve se basear firmemente nas características e qualidades do sujeito. De modo que a personalidade tem um componente intrínseco e orgânico, bem como um componente visível e aparente; e pode ser vinculada a características e qualidades específicas do indivíduo passíveis a descrição e à mensuração objetivas (ALLPORT, 1937 apud HALL; LINDZEY; CAMPBELL, 2007).

Uma outra opção de teoria de personalidade é chamada de globalizante ou de coletânea. Para essa teoria, a personalidade de um indivíduo é delimitada pela enumeração dos principais conceitos que descrevem a pessoa e a combinação desses princípios define a personalidade do sujeito. Outra modalidade de definição, identificada por Allport, destaca a função integrativa ou de organização da personalidade. De acordo com esta definição, a personalidade assume um papel de organização e padrão para as diferentes respostas do indivíduo. A personalidade resulta em ordenamento e congruência aos comportamentos apresentados por um sujeito. Essa função poderá ser um impulso básico, uma necessidade, ou simplesmente o fato de que o indivíduo, como ser coordenado, busca organizar seus comportamentos em vez de operar a partir de reflexos desconexos. Por último, Allport menciona a definição de personalidade que enfatiza a função de ajustamento da personalidade. Neste contexto, a personalidade é constituída por características e comportamentos que permitem ao indivíduo ajustar-se e viver em seu ambiente. A personalidade é delimitada pelas ações que os indivíduos realizam e que os auxiliem a manter equilíbrio, ou que os faça permanecer em harmonia no ambiente que os cercam (HALL; LINDZEY; CAMPBELL,

2007).

A quantidade de alternativas de teorias é extensa, o quê demonstra a falta de consenso entre os teóricos de uma definição universal a cerca da personalidade.

### **2.1.2 Teoria dos Traços da Personalidade**

Uma das teorias mais recentes relacionadas ao estudo da personalidade humana se concentra na tradição psicométrica, que utiliza métodos quantitativos para mensurar e estudar as diferenças individuais (ANDRADE, 2008). Allport, na década de 1930, aplicou a técnicas de análise léxica e análise fatorial no estudo da personalidade, o que contribuiu para a formulação da teoria baseada no conceito de traços. A lógica por trás da abordagem léxica segue a ideia de que a personalidade deve se manifestar na linguagem humana, que usamos diariamente para descrever a nós mesmos e aos outros. Seguindo essa linha de raciocínio e contando com o auxílio de métodos estatísticos, a análise do uso da linguagem oferece fortes indícios sobre os traços de personalidade (ALLPORT; ODBERT, 1936). O teórico considerava os traços de personalidade como um sistema neuropsíquico, intrínseco a cada indivíduo, capaz de receber e converter estímulos funcionalmente equivalentes em padrões de comportamento congruentes. Em outras palavras, os traços de personalidade são predisposições a responder igualmente ou de modo semelhante a determinados estímulos. Segundo Andrade, os traços de personalidade apresentam as seguintes propriedades:

1. Os traços de personalidade possuem um aspecto concreto e que estão presentes em todo indivíduo; eles não são conceitos puramente teóricos ou apenas simbologias concebidas para descrever comportamentos;
2. Os traços provocam ou governam o comportamento; eles são reflexos de certos estímulos e sofrem influência do ambiente para gerar os comportamentos;
3. Os traços estão inter-relacionados e podem sobrepor-se, embora representem características diferentes, e;
4. Os traços, dependendo da circunstância ou cenário, sofrem variações; por exemplo, uma pessoa, numa determinada situação, pode externar o traço de nervosismo, enquanto em um outro contexto pode exibir o traço de serenidade.

Também segundo Andrade, os traços de personalidade são instrumentos úteis para a realização de algumas análises e considerações:

- Os traços podem ser utilizados para resumir e descrever a conduta de um indivíduo;
- Podem ser empregados para realizar previsões sobre o comportamento futuro de uma pessoa;

- Os traços de personalidade preconizam que o padrão comportamental de uma pessoa está associado a características intrínsecas do sujeito; os traços promovem a ideia que um tipo de processo ou mecanismo interno dá origem ao comportamento.

### **2.1.3 Origem e estabilidade da personalidade**

Os traços de personalidade são reflexos de fatores hereditários e ambientais. A hereditariedade fornece a matéria prima da personalidade, que pode ser moldada, ampliada ou limitada pelas condições do ambiente (ANDRADE, 2008). No entanto, não há consenso entre os pesquisadores sobre o quanto cada fator contribui para a definição da personalidade, alguns pesquisadores atribuem maior peso aos fatores hereditários, enquanto que outros entendem que os aspectos ambientes são mais importantes.

Um outro aspecto a ser destacado na teoria do traços de personalidade diz respeito a questão da estabilidade da personalidade em função do tempo e em relação à situação. Em se tratando da estabilidade em relação ao tempo, estudos longitudinais têm demonstrado que a personalidade tende a permanecer estável e inalterada quando o indivíduo atinge a vida adulta (MCCRAE; COSTA, 1994a; MCCRAE; COSTA, 1994b; NETTLE, 2006; EDMONDS et al., 2008). Isto é, uma vez que uma pessoa passa pelas fases da infância e da adolescência, seus traços tendem a não se alterarem durante todo o restante da vida.

### **2.1.4 Modelos de Traços da Personalidade**

Após a formulação da teoria dos traços da personalidade, a próxima questão a ser pesquisada passou a ser quantos traços diferentes seriam necessários para representar a personalidade humana. Utilizando a mesma técnica empregada por Cattell na década de 1930, a técnica de análise fatorial, outros pesquisadores dedicaram-se ao tema e propuseram diferentes modelos para a teoria de traços, sendo que cada modelo proposto é constituído por um número diferente de dimensões de personalidade.

Por exemplo, Cattell em 1943 adotando uma abordagem puramente estatística, desprovido de ideias pré-concebidas, chegou a um modelo composto por 16 fatores ou traços básicos de personalidade. Cattell considera os traços como elementos mentais da personalidade. De acordo com esta concepção, conhecer os traços de um sujeito é pré-requisito essencial para prognosticar como esse indivíduo irá se comportar numa determinada situação.

Allport e Odbert (1936) definiram os traços como unidades básicas da personalidade, que explicam como o indivíduo age regularmente em situações diferentes com o passar do tempo. No modelo proposto por Allport e Odbert os traços de personalidade estão organizados em três dimensões principais:

- Traços cardeias, que representam forças poderosas que dominam o comportamento do sujeito, por exemplo: sadismo e patriotismo fanático. De acordo com o teórico, os indivíduos,

em geral, possuem poucos traços cardeais;

- Traços centrais que compreendem tendências de comportamento num universo menor de situações em comparação com os traços cardeais. Essa dimensão inclui, por exemplo, traços como honestidade, bondade e assertividade;
- Traços secundários que representam disposições de comportamento menos visíveis generalizados e consistentes.

Um terceiro modelo de traços de personalidade, proposto por Eysenck em 1953 e 1983, preconiza os traços de personalidade como representações das intercorrelações de comportamentos específicos e repetidos. Eysenck, no seu modelo, apresenta três dimensões básicas de personalidade: “Introversão – Extroversão”, “Neuroticismo” e “Psicoticismo”. A primeira dimensão engloba traços como, por exemplo, extroversão e humildade; enquanto que a segunda contempla traços como tenso, mal-humorado e baixa auto-estima. Já a terceira dimensão está presente em indivíduos com tendências a apresentar algum tipo de psicopatia, que envolve, por exemplo, traços como impulsividade e crueldade (ANDRADE, 2008).

A seguir é apresentada uma breve discussão sobre uma outra proposta de modelo traço de personalidade: o modelo dos Cinco Grandes Fatores da Personalidade, que servirá como instrumento base para realização da investigação desta dissertação de mestrado.

### **2.1.5 Modelo dos Cinco Grandes Fatores da Personalidade**

O modelo dos Cinco Grandes Fatores (CGF), também conhecido como modelo *Big Five*, alcançou status de modelo de referência e destaque dentre as demais alternativas de modelos de traços de personalidade existentes. As cinco dimensões do modelo foram identificadas por meio de análise das palavras que as pessoas utilizam naturalmente para descrever a si próprios e os demais indivíduos (SRIVASTAVA; JOHN, 1999 apud ANDRADE, 2008).

O modelo foi proposto por Costa e McCrae em 1992a, que o conceberam como uma representação dos traços de personalidade. De acordo com Andrade (2008):

Os cinco grandes fatores devem ser percebidos como tendências básicas que possuem uma base biológica. Essas tendências básicas – disposições para agir e sentir de determinadas maneiras – não são influenciadas diretamente pelo ambiente. Os autores sugerem que existe uma maturação intrínseca da personalidade, segundo a qual os traços de personalidade funcionam mais propriamente como expressões da biologia humana do que como produtos de experiências da vida. Tanto as adaptações características quanto as influências ambientais externas determinam as escolhas e decisões que o indivíduo faz com o passar do tempo, que são refletidas na sua biografia objetiva e também no seu autoconceito.

A justificação de que os CGF representam dimensões básicas da personalidade é baseada em quatro linhas de raciocínio e evidência empírica:

1. Estudos longitudinais e de observação cruzada têm demonstrado que os cinco fatores são disposições duradouras que se manifestam em padrões de comportamentos;
2. Os traços relatados para cada fator são encontrados em uma variedade de teorias de personalidade, bem como na linguagem usual de descrição desses traços;
3. Os fatores são encontrados em diferentes idades, sexos, raças e nacionalidades, embora eles variem em algum grau nas diferentes culturas;
4. Evidências de hereditariedade sugerem que os fatores possuem uma base biológica.

O modelo pressupõe avaliar a personalidade humana sob cinco dimensões, descritas a seguir:

**Abertura (Openness):** Indivíduos com alta pontuação nesta dimensão, geralmente, são francos, imaginativos, espirituosos, originais e artísticos. Por outro lado, indivíduos com baixa pontuação nesta dimensão são superficiais, comuns ou simples.

**Conscienciosidade (Conscientiousness):** Nesse fator é característico o controle de impulsos, bem como comportamentos direcionados a um objetivo específico, que podem facilitar a execução de obrigações e deveres. Indivíduos conscienciosos são geralmente cautelosos, dignos de confiança, organizados e responsáveis. Por outro lado, indivíduos com baixos escores nessa dimensão tendem a ser descuidados, desordenados e pouco-confiáveis.

**Extroversão (Extraversion):** Também é chamada de “Expansão”; indivíduos extrovertidos tendem a ser ativos, entusiasmados, dominantes, sociáveis e eloquentes ou falantes. Por outro lado, indivíduos introvertidos tendem a ser retraídos, submissos e quietos.

**Amabilidade (Agreeableness):** Também é chamado comumente de “Agradabilidade” ou “Sociabilidade”. Indivíduos com altas pontuações nesse traço são agradáveis, amáveis, cooperativos e afetuosos. Indivíduos com uma classificação baixa nessa dimensão podem ser frios e indelicados. A referida dimensão caracteriza-se por uma orientação em direção aos demais, incluindo traços como altruísmo, confiança e modéstia.

**Neuroticismo (Neuroticism):** Também é chamado de “Instabilidade emocional”. Indivíduos neuróticos são geralmente nervosos, altamente sensíveis, tensos e preocupados. Por outro lado, indivíduos emocionalmente estáveis são calmos e satisfeitos. O traço Neuroticismo contrasta estabilidade emocional com afetos negativos, incluindo ansiedade, tristeza, irritabilidade e tensão nervosa.



A estrutura penta-dimensional do instrumento foi validada e se mostrou robusta na maioria das culturas e regiões do mundo (MCCRAE et al., 1998; ROLLAND, 2002; SCHMITT et al., 2007) (ROLLAND, 2002 apud ANDRADE, 2008), tendo isso também comprovada sua adequação à realidade brasileira (ANDRADE, 2008).

#### **2.1.5.1 Inventário dos Cinco Grandes Fatores da Personalidade (IGFP-5) – *The Big Five Inventory***

Existem diversas opções de instrumentos elaborados com o propósito de medir os Cinco Grandes Fatores da Personalidade, incluindo um dos mais divulgados, o instrumento *NEO Personality Inventory Revised (NEO-PI-R)* concebido por Costa e McCrae (1992b) e formado por 240 itens. O preenchimento deste questionário é concluído, em média, em 45 minutos; o quê acaba por reduzir sua aplicabilidade em estudos de carácter científico (PALMA, 2006), principalmente em pesquisas que exigem amostras com números maiores de participantes. Devido a esta limitação, foram sugeridas alternativas de instrumentos que apresentassem um número menor de itens. Dentre as alternativas existentes, destacam-se os instrumentos: *Big Five Inventory* (BFI; John, Donahue e Kentle (1991a)) e *NEO Five-Factor Inventory* (NEO-FFI; Costa e McCrae (1992b)).

Srivastava e John (1999) estimaram que o BFI (IGFP-5) e NEO-FFI levam aproximadamente cinco e dez minutos para serem preenchidos. Mesmo possuindo um número menor de itens e exigir menos tempo para ser concluído, Benet e John indicam que os instrumentos BFI (IGFP-5) e NEO-FFI apresentam resultados equivalentes, por essa razão o Inventário dos Cinco Grandes Fatores da Personalidade (IGFP-5) foi adotado nesta pesquisa.

O instrumento de questionário usado nesta dissertação foi desenvolvido em 1991 pelo Instituto de Personalidade e Pesquisa Social da Universidade de Berkeley num projeto de pesquisa liderado por Oliver P. John (JOHN; DONAHUE; KENTLE, 1991a; JOHN; DONAHUE; KENTLE, 1991b). O instrumento foi posteriormente adaptado e validado para diferentes línguas (BENET; JOHN, 1998; SCHMITT et al., 2007; DENISSEN et al., 2008; ZHENG et al., 2008; FOSSATI et al., 2011; ZAMORANO et al., 2014), inclusive para o português (ANDRADE, 2008; BRITO-COSTA et al., 2016). O inventário permite uma avaliação eficiente, flexível, de mensuração rápida e diferenciada dos cinco grandes fatores: Abertura à Experiência, Amabilidade, Conscienciosidade, Neuroticismo (Estabilidade Emocional) e Extroversão (SRIVASTAVA; JOHN, 1999).

O instrumento está disponível para ser usado gratuitamente em pesquisas científicas e para outros fins não comerciais.

O questionário é formado por 44 itens, representados por sentenças curtas e que as respostas seguem a escala *Likert* de cinco pontos, variando entre (1) "Discordo totalmente" a (5) "Concordo totalmente". A versão do instrumento IGFP-5 traduzida para o português pode ser visualizada no Anexo A deste trabalho.

### **2.1.6 O impacto das diferenças individuais na aprendizagem**

A relação entre os cinco fatores da personalidade e o desempenho na aprendizagem tem sido amplamente pesquisada (CASPI et al., 2006; RICHARDSON; ABRAHAM; BOND, 2012). Diversos estudos têm se dedicado a investigar o efeito da personalidade no aprendizado (RAAD; SCHOUWENBURG, 1996), tendo sido possível identificar correlações entre os cinco traços de personalidade e vários temas relacionados ao contexto da aprendizagem.

Traços de conscienciosidade estão presentes em indivíduos que tendem a ser confiáveis, organizados, persistentes e orientados a objetivos. Aqueles com alto nível de conscienciosidade são mais motivados em realizar suas atividades com afinco e são mais persistentes quando confrontados com dificuldades e desafios. Pesquisas têm demonstrado a existência da correlação direta entre o traço de conscienciosidade e o desempenho acadêmico (CONARD, 2006; POROPAT, 2009; VEDEL, 2014).

Estudantes com traços de abertura para experiência são mais imaginativos, criativos e dispostos a considerar novas ideias. Esses alunos demonstram ser mais capazes em gerir novas aprendizagens essenciais para o desempenho acadêmico (VERMETTEN; LODEWIJKS; VERMUNT, 2001). Segundo Lounsbury et al. (2003), alunos com escores maiores para os traços de abertura e amabilidade são mais inclinados em comparecer às atividades pedagógicas com mais consistência, apresentando índices de ausências menores. Enquanto que alunos com traços de amabilidade expressam mais facilidade em firmar cooperação com os instrutores, tornando o processo de aprendizagem mais fácil (VERMETTEN; LODEWIJKS; VERMUNT, 2001).

O traço de personalidade de neuroticismo está associado a altos níveis de ansiedade (STEEL; BROTHEN; WAMBACH, 2001), que acabam por comprometer o desempenho de alunos em testes e exames.

Segundo Richardson, Abraham e Bond (2012), o traço de extroversão implica em maior sociabilidade. Segundo Rolfhus e Ackerman (1999), pode-se esperar que alunos com tendências extrovertidas alcancem notas mais baixas, porque são mais distraídos e mais sociáveis do que estudantes com tendências introvertidas, que tendem a gastar mais tempo aprendendo e consolidando o conhecimento. De acordo com Bidjerano e Dai (2007), o traço de extroversão pode limitar a capacidade dos alunos de regular seu esforço dedicado a tarefas acadêmicas.

## **2.2 Mineração de Dados Educacionais**

A Mineração de Dados Educacionais (do inglês, Educational Data Mining – EDM) é a área de pesquisa de natureza interdisciplinar que estuda o desenvolvimento de métodos e algoritmos de análise de informações relacionadas ao meio educacional (ROMERO; VENTURA, 2010). Ela aplica uma abordagem estatística e computacional para analisar grandes massas de dados, com o intuito de explorar e oferecer respostas pertinentes no contexto educacional.

As primeiras iniciativas de uso de técnicas de mineração de dados em problemas da área

educacional remetem à primeira metade da década de 2000, com pesquisas que buscaram estudar a aplicação de métodos e técnicas de mineração de dados para o diagnóstico da aprendizagem em ambientes de ensino a distância (ROMERO; VENTURA, 2007).

No trabalho de Romero e Ventura (2010), várias aplicações de técnicas de mineração de dados no contexto educacional são listadas, dentre as quais:

**Modelagem de alunos:** o objetivo dessa classe de atividade é identificar determinados atributos dos alunos. Essa aplicação está relacionada ao desenvolvimento de modelos que podem ser usados para representar os alunos e identificar algumas das suas características, como por exemplo: motivação (COCEA; WEIBELZAHN, 2011), satisfação (LEE, 2010), desempenho de aprendizagem (KOTSIANTIS; PIERRAKEAS; PINTELAS, 2004; ROMERO et al., 2013) e status emocional (BAKER et al., 2010).

**Prover recomendações:** tem como meta identificar e recomendar aos alunos conteúdos pedagógicos que sejam mais apropriados para os alunos, levando-se em conta suas necessidades atuais e seus níveis de aprendizagem.

**Prever o desempenho de alunos:** busca prever o resultado futuro de testes e avaliações pedagógicas, com as previsões sendo baseadas no histórico das atividades realizadas pelo aluno.

**Comunicação dos envolvidos:** tem como propósito tornar mais proveitosa e eficiente a participação dos alunos nas atividades, bem como otimizar a comunicação entre administradores, equipes pedagógicas e professores integrantes dos cursos.

**Facilitar melhorias em cursos:** objetiva auxiliar as equipes pedagógicas e professores a realizar implantar estratégias que tragam melhorias para aprendizagem dos alunos.



### 3 TRABALHOS RELACIONADOS

Neste capítulo, são apresentados alguns trabalhos relacionados ao tema desta pesquisa. Foram considerados trabalhos que tratam sobre a classificação da personalidade de estudantes no contexto do ensino à distância. Para cada artigo, buscou-se detalhar seus objetivos, a metodologia de pesquisa aplicada, os resultados obtidos, suas contribuições e suas possíveis limitações. Em particular, foram abordados os trabalhos de Halawa, Shehab e Hamed (2015), Chen et al. (2016) e Ghorbani e Montazer (2015).

Este capítulo se encerra com a apresentação de uma tabela comparativa que resume as principais características dos trabalhos analisados.

#### 3.1 Predicting Student Personality Based on a Data-Driven Model from Student Behavior on LMS and Social Networks

O objetivo da pesquisa de Halawa, Shehab e Hamed (2015) foi propor um modelo para identificar o tipo de personalidade dos alunos, segundo a classificação tipológica de Myers-Briggs. A exemplo do modelo dos cinco grandes fatores, a classificação Myers-Briggs é um instrumento utilizado para identificar características e preferências pessoais; no entanto, este modelo não se baseia na teoria dos traços de personalidade; segue um arcabouço teórico diferente, assentado na teoria analítica proposta por Carl Jung. O instrumento Myers-Briggs tem como ideia central que as pessoas tendem a considerar que certas maneiras de pensar e agir são mais fáceis ou mais apropriadas do que outras e essas escolhas acabam por definir suas personalidades. O modelo Myres-Briggs considera quatro dimensões organizadas sob a forma de pares opostos de preferências de pensar e agir (Tabela 1).

Tabela 1 – As dimensões (preferências) e as definições do modelo de personalidade classificação tipológica de Myers-Briggs

<b>Preferências</b>	<b>Definições</b>
Extroversão (Extroversion) ou Introversão (Introversion)	Onde o indivíduo obtém sua energia
Sensoriais (Sensing) ou Intuitivas (Intuition)	A maneira como um indivíduo obtém uma informação
Racionalistas (Thinking) ou Sentimentais (Feeling)	A maneira como um indivíduo toma decisões
Julgadores (Judging) ou Perceptivos (Perceiving)	Como o indivíduo lida com o ambiente externo

Adaptado da obra de Halawa, Shehab e Hamed (2015).

A combinação das quatro preferências estabelece dezesseis tipos diferentes de personalidades, que são representadas pelas iniciais das dimensões mais significativas para o indivíduo.

Para a elaboração do modelo, os pesquisadores utilizam os logs das interações dos estudantes registradas no Ambiente Virtual de Aprendizagem e dados de uma rede social, neste caso o Facebook. A pesquisa foi realizada com uma amostra de 240 estudantes da faculdade de negócios da Universidade Alemã do Cairo, Egito.

Inicialmente, o questionário do instrumento foi aplicado aos estudantes da amostra. As respostas obtidas foram organizadas e analisadas, e o tipo de personalidade de cada estudante foi estabelecido em conformidade com as regras do instrumento. No estágio seguinte, os pesquisadores utilizaram os resultados do questionário, alguns atributos gerados por meio dos logs do ambiente de aprendizagem da instituição de ensino (Tabela 2) e alguns outros atributos da rede social Facebook (Tabela 3) como entradas para o treinamento de algoritmos de aprendizagem de máquina supervisionada, que elaboraram modelos preditores de personalidade. Esses modelos foram validados e seus resultados foram comparados. Os algoritmos usados no experimento foram: NaiveBayes, BayesNet, Kstar, Floresta Aleatória (*Random Forest*), J48, OneR, JRIP, KNN/IBK, RandomTree, e Tabela de Decisão (*Decision Table*).

Tabela 2 – Atributos dos alunos definidos pela pesquisa de Halawa, Shehab e Hamed (2015) e recuperados a partir do ambiente de aprendizagem da Universidade Alemã do Cairo

<b>Atributo</b>	<b>Descrição</b>
Núm. de páginas visitadas	Número de páginas visitadas no sistema por alunos durante todo o semestre acadêmico
Total de tempo no ambiente	Tempo gasto por alunos em páginas, chat e em outras atividades do ambiente de aprendizagem
Núm. de comentários	Número de comentários escritos pelos alunos nas diferentes páginas do curso
Núm. de posts	Número de posts realizados nas atividades do tipo blog
Núm. de sessões de chat	Número de participações em atividades de chat
Núm. de tarefas entregues antes do prazo	Número de tarefas resolvidas e entregues antes do prazo estipulado pelo instrutor do curso
Núm. de tarefas entregues com atraso	Número de tarefas entregues com atraso em relação ao prazo estipulado pelo instrutor do curso

Tabela 3 – Atributos dos alunos relacionados às interações da rede social escolhida pela pesquisa de Halawa, Shehab e Hamed (2015)

<b>Atributo</b>	<b>Descrição</b>
Núm. de likes	Número de likes na rede social realizados por alunos sobre as páginas dos cursos
Núm. de compartilhamentos	Número de compartilhamentos efetuados na rede social realizados por alunos sobre as páginas dos cursos

De acordo com os autores do artigo, os modelos construídos permitem predições com um alto grau de acurácia, pelo menos seis dos algoritmos avaliados alcançaram percentuais de acerto maiores que 85%; destacando-se o algoritmo OneR, que segundo o texto, atingiu percentual 97% de acurácia.

Apesar dos resultados obtidos nos experimentos terem sido positivos, há necessidade de realizar algumas considerações sobre as escolhas feitas pelos pesquisadores para a análise das informações. Para a fase de construção dos modelos, os pesquisadores decidiram dividir a base de dados em dois grupos, na proporção: 20% das instâncias foi alocada para a etapa de treinamento dos algoritmos de aprendizagem, enquanto 80% foi reservado para a fase de validação. Isto é, para a construção do modelo foram utilizados os atributos de apenas 48 alunos. No geral, as obras de referência de mineração de dados e aprendizagem de máquina recomendam uma divisão diferente, algo em torno de 70% – 80% para o grupo de treinamento e o restante para os grupos de validação e testes (GOLLAPUDI, 2016). Geralmente, quanto maior a amostra de treinamento, melhor será o modelo preditor (WITTEN; FRANK; HALL, 2011); sendo necessário tomar algumas precauções, durante a definição dos dados de treinamento, com o intuito de evitar que o modelo construído se torne demasiadamente ajustado ao conjunto de dados de treinamento e de validação, o que o impossibilitará de realizar predições úteis e mais próximas da realidade (GOLLAPUDI, 2016). Quando esse tipo situação indesejada ocorre, diz-se que o modelo passou por processo de sobreajuste (do inglês, *overfitting*).

Ainda sobre a análise da metodologia, adotada na pesquisa, e com o objetivo de reduzir o viés na seleção de dados destinados ao treinamento e à validação do modelo; poderia ter sido sondada pelos autores da pesquisa a possibilidade de usar uma das técnicas de validação cruzada (*cross-validation*) com o propósito de tornar a avaliação do modelo mais robusta.

### **3.2 On the Impact of Personality in Massive Open Online Learning**

O estudo de Chen et al. (2016) tem como meta investigar o impacto da personalidade dos alunos no ensino à distância em ambientes educacionais do tipo *Massive Open Online Courses* (MOOCs). A investigação busca responder duas questões:

**Pergunta 1:** No contexto de plataformas MOOCs, o tipo de personalidade gera algum impacto no engajamento, comportamento e no sucesso do aluno?

**Pergunta 2:** A personalidade dos alunos pode ser identificada ao analisar suas ações e comportamentos na plataforma MOOC?

A pesquisa foi realizada num curso de curta duração, intitulado *Data Analysis*, ofertado na plataforma de ensino edX ([www.edx.org](http://www.edx.org)) entre os meses de Agosto e Novembro de 2015. A princípio, o curso registrou a inscrição de 23.622 usuários, mas apenas 40% deles realizaram pelo

menos uma das atividades do plano do curso e a taxa de conclusão foi de 4,75% dos inscritos, percentual similar a de outros cursos MOOCs (KOLLER et al., 2013).

No início do curso, os pesquisadores solicitaram aos participantes que respondessem um questionário de classificação de personalidade. Os responsáveis pela pesquisa optaram pelo instrumento que segue o modelo dos cinco grande fatores de personalidade. Dos usuários que se inscreveram e que participaram de pelo menos uma das atividades do curso, 763 usuários responderam o questionário integralmente e tiveram seus dados analisados na pesquisa.

Além dos resultados do questionário de personalidade, a pesquisa utilizou vinte atributos adquiridos a partir dos logs das interações dos participantes. A tabela 4 exhibe alguns dos atributos adquiridos a partir do log de interações da plataforma edX.

Tabela 4 – Atributos dos alunos recuperados por meio das interações com o ambiente de ensino edX durante as atividades do curso *Data Analysis* e utilizados na pesquisa de Chen et al. (2016)

<b>Atributo</b>	<b>Descrição</b>
Tempo dedicado para assistir vídeos do conteúdo	Total de tempo dedicado assistir os vídeos com o conteúdo do curso
Tempo resolvendo <i>quizzes</i>	Tempo gasto pelo aluno em páginas do tipo quiz
Núm. de questões respondidas	Total de questões que os alunos tentaram responder (independentemente de certos ou erros)
Núm. de posts em fóruns	Número de novos posts criados em fóruns
Tempo dedicado visitando fóruns	Total de tempo dedicado para visualizar os fóruns do curso
Tempo total na plataforma	Quantidade de tempo (em minutos) que aluno permaneceu na plataforma de ensino
Núm. de vídeos ignorados	Número de vídeos ignorados e não assistidos
Tempo médio das sessões	Tempo médio das sessões na plataforma edX
Tempo médio para responder perguntas do mesmo quiz	Tempo médio entre as respostas das perguntas do mesmo quiz

Para investigar a primeira questão, os pesquisadores empreenderam uma análise correlacional entre as interações dos alunos na plataforma edX e os traços de personalidade diagnosticados pelo questionário. Na análise, os pesquisadores optaram em utilizar o coeficiente de correlação de Spearman, e todos os coeficientes calculados forneceram valores próximos de zero, sinalizando que não foi possível estabelecer uma correção estatisticamente significativa entre os traços de personalidade e os comportamentos dos alunos na plataforma edX.

Em relação à segunda questão, o trabalho utilizou algoritmos de aprendizagem supervisionada para estabelecer modelos computacionais que pudessem classificar o tipo de personalidade dos alunos através da análise estatística dos padrões das interações dos alunos com o ambiente



MOOC edX. O artigo afirma que os algoritmos de classificação Floresta Aleatória (*Random Forest*) e Processo Gaussiano (*Gaussian Process*) foram selecionados para a análise. Para a construção e validação dos modelos, os pesquisadores optaram em aplicar o método de validação cruzada *k-fold*, parametrizado para dividir a amostra dos dados em 10 grupos. Ambos algoritmos geraram modelos que se mostraram ineficientes em gerar previsões acuradas. Os melhores resultados de previsão estão associados ao traço de personalidade de extroversão, atingindo coeficiente Spearman próximo a 0,30. Os modelos obtiveram coeficientes de correlação ainda mais baixos para as outras dimensões de personalidade.

Infelizmente, para a análise dos dados, a pesquisa levou em consideração as atividades de um único curso. É provável que modelos mais precisos e robustos pudessem ter sido gerados se a pesquisa também contemplasse estudantes de outros cursos.

### 3.3 E-learners' personality identifying using their network behaviors

Os pesquisadores Ghorbani e Montazer (2015) conduziram um experimento num ambiente real de ensino à distância com a proposta de desenvolver um modelo computacional capaz de prever os traços de personalidade dos alunos participantes. Os dados analisados foram adquiridos de um curso de graduação on-line ministrado na Universidade Tarbiat Modares com sede em Teerã, no Irã. A pesquisa abordou duas questões relacionadas à personalidade dos alunos de ensino à distância:

**Pergunta 1:** Quais comportamentos são relacionados às dimensões de personalidade?

**Pergunta 2:** Quais são as relações entre os comportamentos de rede e as dimensões de personalidade?

Para responder à primeira questão, os idealizadores da pesquisa aplicaram uma versão do inventário dos cinco grandes fatores de personalidade (inventário NEO-FFI), desenvolvido por Costa e McCrae (1992b), com 60 itens para avaliar as dimensões de Conscienciosidade, Neuroticismo, Extroversão, Amabilidade e Abertura. O questionário foi aplicado para uma amostra de 95 alunos. Além da aplicação do questionário, o estudo também considerou dez atributos gerados a partir dos logs do ambiente de aprendizagem (Tabela 5). Testes estatísticos foram utilizados com o propósito de confirmar a existência de correlações entre comportamentos apresentados pelos estudantes no ambiente virtual de aprendizagem e os resultados do inventário de personalidade.

O experimento identificou correlações estatisticamente significativas nas cinco dimensões do modelo de personalidade. Por exemplo, o estudo aponta que o atributo “Número de visualizações dos avisos das lições” demonstra correlação direta com o traço de personalidade conscienciosidade, com coeficiente Spearman de 0,85.

Sobre a segunda questão, foi implementado um outro experimento no qual o mesmo questionário foi respondido por uma amostra de 53 estudantes que tiveram os mesmos atributos (Tabela 5) analisados por um algoritmo de classificação, baseado em lógica difusa, para definir um modelo computacional que permita o prognóstico do tipo de personalidade de participantes do curso. Os pesquisadores optaram em utilizar a técnica de validação cruzada para construir e avaliar o modelo implementado. O estudo mostrou que o modelo conseguiu gerar previsões precisas sobre as personalidades dos alunos. Segundo os autores, o modelo conseguiu atingir acurácias entre 78% a 97% de acerto nas previsões para as cinco dimensões do modelo de personalidade.

O modelo desenvolvido proporcionou resultados com níveis de precisão altos, mas ele foi desenvolvido levando em consideração uma amostra de participantes relativamente pequena, faz-se necessário construir novas versões do modelo que se baseiem em um número maior de casos e que considerem cursos de diferentes áreas do conhecimento.

Tabela 5 – Atributos dos alunos relacionados às interações no ambiente virtual de aprendizagem escolhidos pela pesquisa de Ghorbani e Montazer (2015)

<b>Atributos</b>
Núm. de amigos
Pontuação final
Participação em chats e fóruns
Núm. de posts em fóruns
Nível de dificuldade dos exercícios e exemplos
Tempo dedicado para leitura de materiais do curso
Tempo dedicado para leitura de exercícios e exemplos do curso
Participação em atividades de soluções de problemas em grupo
Atrasos (em dias) na entrega de tarefas do curso
Núm. de entradas no sistema por semana
Visualizações de novidades sobre as lições

### 3.4 Tabela Comparativa

De modo a facilitar o entendimento de trabalhos relacionados com o trabalho apresentado nesta dissertação, a tabela 6 apresenta um resumo esquemático comparativo. Neste resumo estão listados os trabalhos relacionados segundo suas propostas de abordar ou não a análise de informações em Ambientes Virtuais de Aprendizagem (AVA), de apresentar um algoritmo para classificar a personalidade de forma automática, de adotar o instrumento dos cinco grandes fatores como modelo psicológico basilar para a realização da pesquisa e, por fim, se o modelo construído passou por um processo de validação.

Tabela 6 – Tabela comparativa dos trabalhos relacionados

	<b>Halawa, Shehab e Hamed, 2015</b>	<b>Chen, Davis e Houbem, 2016</b>	<b>Ghorbani e Montazer, 2015</b>
<b>Aborda Sistemas de Gestão de Aprendizagem</b>	Sim	Não*	Sim
<b>Considera dados de cursos de diferentes áreas?</b>	Não	Não	Não
<b>Apresenta um modelo para classificação automática de personalidade?</b>	Sim	Sim	Sim**
<b>Adota o modelo <i>Big Five</i> para classificar a personalidade?</b>	Não	Sim	Sim
<b>Apresenta processo de validação?</b>	Sim	Não	Sim

\* Aborda MOOCs

\*\* O modelo foi treinado com auxílio de uma base de dados com poucos registros



## 4 PROPOSTA DA PESQUISA

Os traços de personalidade estabelecem padrões comportamentais que as pessoas manifestam continuamente e rotineiramente nas diversas atividades do dia a dia, incluindo, inclusive, a prática da aprendizagem. Diversas pesquisas têm demonstrado a relevância e aplicabilidade do conceito de personalidade para o contexto educacional e também têm demonstrado que indivíduos com traços de personalidade diferentes tendem a se comportar e perceber o processo aprendizagem de maneiras distintas (RAAD; SCHOUWENBURG, 1996). Assuntos como desempenho ou performance da aprendizagem, persistência, motivação, auto-regulação e orientação aos estudos estão relacionados com o conceito de personalidade.

Como exemplo do que foi mencionado no parágrafo anterior, pesquisas têm demonstrado uma correlação direta entre os traços de personalidade de alunos e seus desempenhos em avaliações e exames dos cursos. De acordo com esses estudos, alunos que possuem certos traços de personalidade tendem a alcançar melhores pontuações do que indivíduos com outros estilos de personalidade (POROPAT, 2009; VEDEL, 2014; ROLFHUS; ACKERMAN, 1999; BUSATO et al., 2000; FURNHAM; MONSEN, 2009). Outros estudos correlacionam a concepção de personalidade a temas como motivação para aprender (ARIANI, 2013), persistência na realização das atividades acadêmicas (FALCONER; ADRAGNA, 2017; SATCHELL et al., 2017) e auto-regulação e orientação nos estudos (FRUYT; MERVIELDE, 1996; FABIO, 2006).

Dada a influência do constructo da personalidade em vários aspectos da aprendizagem, torna-se latente a necessidade de oferecer recursos e ferramentas que possibilitem a identificação dos traços de personalidade dos alunos. Ao conhecer o estilo de personalidade dos alunos, abre-se a possibilidade para que docentes, e demais membros da equipe pedagógica, antevejam algumas situações e realizem intervenções na estratégia pedagógica com o objetivo de tornar a aprendizagem mais eficiente e personalizada.

No contexto do ensino a distância, onde, em geral, os docentes e a equipe pedagógica possuem poucas oportunidades de contato direto com os alunos, a maneira mais viável de identificar os traços de personalidade dos discentes se dá por meio da aplicação de questionários. Tradicionalmente, o diagnóstico da personalidade é realizado através da aplicação de instrumentos psicométricos (ANDRADE, 2008). Entretanto, o uso rotineiro deste recurso demanda muito esforço; tanto para aplicá-lo, quanto para quantificar e interpretar seus resultados. É muito provável que a equipe pedagógica não disporá de recursos e tempo suficientes para realizar e gerir manualmente todo esse processo. Além disso, o uso de questionários apresenta algumas limitações e restrições, por exemplo:

- Existe a possibilidade de solicitar o preenchimento de questionários por meio eletrônico, mas, em geral, o percentual de preenchimento e retorno completo para esse tipo de contato é baixo;

- Há margem para o preenchimento errado (deliberado ou não) dos questionários, possibilitando a ocorrência de respostas influenciadas pelo interesse em ampliar ou omitir características do respondente.

Então, considerando a problemática e as restrições relativas ao uso de questionários, podemos perceber que há justificativa para pesquisar uma alternativa a esta abordagem. Sendo assim, esta pesquisa pretende investigar a viabilidade de construir um modelo computacional que possibilite a predição automática do tipo de personalidade dos alunos por meio de técnicas de mineração de dados educacionais.

Segundo esta proposta, a investigação por um modelo será baseada exclusivamente nos registros dos hábitos, ações e interações realizados pelos alunos durante as atividades pedagógicas realizadas em um Sistema Gestor de Aprendizagem (no caso, o sistema Moodle<sup>1</sup>). E que, caso seja viável, este modelo possa servir de ferramenta prática para que professores, tutores e equipe pedagógica possam identificar os traços de personalidade dos alunos, dispensando totalmente a necessidade de aplicação de questionários psicométricos.

O projeto desta pesquisa, seus procedimentos e instrumentos (questionário) foram submetidos ao Comitê de Ética (CPE) da Universidade Federal de Alagoas no dia 12/09/2018, e foi aprovado e autorizado no dia 13/12/2018. O parecer final do CEP consta no **Anexo C**.

## **4.1 Construção da base de dados**

Esta seção descreve as etapas para a definição e construção da base de dados que foi utilizada nesta pesquisa. A subseção 4.1.1 apresenta a seleção da amostra, a subseção 4.1.2 apresenta o processo de seleção e aplicação do instrumento tradicional de classificação da motivação. O processo de coleta de logs é apresentado na seção 4.1.3 e, por fim, a seção 4.1.4 apresenta como foi construído o dataset final.

### **4.1.1 Seleção da Amostra**

A realização desta pesquisa foi realizada após autorização da Coordenadoria Institucional de Educação a Distância (CIED), setor da UFAL responsável pela gestão dos cursos EAD da UFAL B, e após a emissão do parecer favorável à pesquisa emitido pelo Comitê de Ética em pesquisa da Universidade (Anexo C).

Após a autorização da CIED, foi definido que a pesquisa teria como alvo os alunos de todos os cursos de graduação à distância ofertados pela Universidade. Um convite de participação foi enviado eletronicamente para os potenciais participantes do experimento, onde a natureza, a extensão e os objetivos do estudo foram apresentados. Cento e treze (113) alunos aceitaram em participar desse estudo. No entanto, cinco desses alunos não foram considerados na

---

<sup>1</sup> [www.moodle.org](http://www.moodle.org)

análise do experimento, pois contabilizaram menos de vinte dias de interação com o ambiente de aprendizagem. Então, a amostra de participantes da pesquisa conta com 108 alunos, sendo 45 do sexo masculino (44%) e 63 do sexo feminino (56%) dos cursos de Administração Pública, Ciências Sociais, Física, Geografia, Letras, Matemática, Pedagogia, Química e Sistemas de Informação. A tabela 7 detalha a distribuição dos participantes do experimento, de acordo com o curso e o gênero dos alunos.

Tabela 7 – Distribuição de alunos participantes do experimento por curso e gênero

Curso	Masculino	Feminino	Total
Administração Pública	1	2	3
Ciências Sociais	2	2	4
Física	3	2	5
Geografia	8	7	15
Letras	3	5	8
Matemática	6	3	9
Pedagogia	1	36	37
Química	2	2	4
Sistemas de Informação	19	4	23
<b>TOTAL</b>	<b>45</b>	<b>63</b>	<b>108</b>

Os dados dos estudantes participantes foram obtidos em duas etapas. Num primeiro momento, um questionário de classificação de personalidade foi respondido pelos alunos, e num segundo estágio, foi coletado um conjunto de registros (*logs*) das interações efetuadas pelos estudantes respondentes no ambiente virtual de aprendizagem utilizado pela instituição de ensino. Estas duas etapas, bem como a construção da coleção de dados final, serão descritas com mais detalhes nas subseções a seguir.

#### 4.1.2 Escolha do Instrumento de Avaliação de Personalidade

Foi aplicado um questionário de personalidade a uma amostra de alunos com o objetivo de se obter uma massa de dados para realizar a pesquisa. Os participantes deste experimento são alunos de cursos de graduação à distância oferecidos pela UFAL<sup>2</sup>.

Em relação ao questionário, esta pesquisa adotou o modelo de questionário *Big Five Inventory* (Inventário dos Cinco Grandes Fatores de Personalidade – IGFP-5), por ser um instrumento que foi adaptado e validado para o Brasil. O IGFP-5 é uma medida de auto-relato breve, composta por 44 itens e designada a avaliar dimensões da personalidade baseada no modelo dos Cinco Grandes Fatores da Personalidade (ANDRADE, 2008), conforme o **Anexo A**. O instrumento IGFP-5 foi escolhido porque demonstrou resultados consistentes em vários estudos anteriores (BENET; JOHN, 1998; SCHMITT et al., 2007; ANDRADE, 2008) e porque apresenta nível de confiabilidade equivalente a outros instrumentos da mesma classe (ver a

<sup>2</sup> Universidade Federal de Alagoas

seção 2.1.5.1), mesmo apresentando um número menor de itens (questões) aos respondentes (SRIVASTAVA; JOHN, 1999).

Para realizar a classificação do estilo de personalidade dos participantes do experimento, as respostas do questionário são aplicadas às equações de pontuação, definidas pelo instrumento de pesquisa. Sendo R a resposta de cada questão e Pt(O) a pontuação do traço de Abertura (*Openness*), Pt(C) a pontuação de Conscienciosidade (*Conscientiousness*), Pt(E) a pontuação de Extroversão (*Extraversion*), Pt(A) a pontuação Amabilidade (*Agreeableness*) e Pt(N) para a pontuação de Neuroticismo (*Neuroticism*), as pontuações dos traços de personalidade são calculadas pelas Equações 4.1, 4.2, 4.3, 4.4 e 4.5.

$$Pt(O) = \frac{R_5 + R_{10} + R_{15} + R_{20} + R_{25} + R_{30} + (6 - R_{35}) + R_{40} + (6 - R_{41}) + R_{44}}{10} \quad (4.1)$$

$$Pt(C) = \frac{R_3 + (6 - R_8) + R_{13} + (6 - R_{18}) + (6 - R_{23}) + R_{28} + R_{33} + R_{38} + (6 - R_{43})}{9} \quad (4.2)$$

$$Pt(E) = \frac{R_1 + (6 - R_6) + R_{11} + R_{16} + (6 - R_{21}) + R_{26} + (6 - R_{31}) + R_{36}}{8} \quad (4.3)$$

$$Pt(A) = \frac{(6 - R_2) + R_7 + (6 - R_{12}) + R_{17} + R_{22} + (6 - R_{27}) + R_{32} + (6 - R_{37}) + R_{42}}{9} \quad (4.4)$$

$$Pt(N) = \frac{R_4 + (6 - R_9) + R_{14} + R_{19} + (6 - R_{24}) + R_{29} + (6 - R_{34}) + R_{39}}{8} \quad (4.5)$$

A estrutura penta-dimensional do instrumento foi validada e se mostrou robusta na maioria das culturas e regiões do mundo (SCHMITT et al., 2007) (ROLLAND2002, 2002 apud ANDRADE, 2008), tendo sido comprovado para a realidade brasileira (ANDRADE, 2008).

A Tabela 8 mostra os valores dos mínimos e máximos, médias e desvios padrão para as pontuações obtidas pelos participantes de experimento a partir das respostas fornecidas no questionário aplicado.

#### 4.1.3 Coleta de Registros (*Logs*) das Interações no Ambiente Virtual de Aprendizagem

Após recepção e tratamento das respostas dos questionários, os dados dos *logs* das interações no ambiente de aprendizagem dos alunos que haviam participado da etapa anterior foram obtidos e passaram por processo de pré-processamento.



Tabela 8 – Caracterização descritiva das pontuações dos traços de personalidade dos participantes do experimento

<b>Traço de Personalidade</b>	<b>Mínimo</b>	<b>Máximo</b>	<b>Média ± Desvio Padrão</b>
Abertura	2,3	4,1	3,231 ± 0,365
Conscienciosidade	1,333	3,667	2,711 ± 0,396
Extroversão	1,625	3,75	2,907 ± 0,385
Amabilidade	1,778	3,444	2,553 ± 0,373
Neuroticismo	2,25	4,375	3,353 ± 0,399

A coleta foi realizada diretamente a partir da base de dados do ambiente virtual de aprendizagem. O acesso aos dados foi autorizado pela Coordenadoria Institucional de Educação a Distância (CIED) da UFAL, responsável pela gestão de cursos de graduação a distância ofertados pela instituição (**Anexo B**).

#### 4.1.4 Construção do conjunto de dados do Experimento

Nesta etapa os dados foram organizados e pré-processados para posterior análise. Os alunos que responderam ao questionário, mas que não participaram de um número mínimo de atividades no ambiente de aprendizagem tiveram suas informações excluídas da coleção de dados a ser utilizada na realização do experimento. As atividades realizadas pelos alunos foram agrupadas e sumarizadas, de tal forma que, para cada discente participante, a base de dados contenha apenas os quantitativos e as médias das interações realizadas no ambiente de aprendizagem, com os números agrupados por aluno e tipo de interação realizada.

Também na etapa de pré-processamento, as respostas e as pontuações, auferidas após a aplicação dos questionários de personalidade, são combinadas com os registros das interações executadas no ambiente de aprendizagem. Isto viabiliza que sejam realizadas tentativas para estabelecer correlações entre os resultados dos questionários de personalidade aplicados e os *logs* de interação dos alunos no ambiente de aprendizagem.

Neste ponto, o conjunto de dados passou por um processo de anonimização, tornando inviável identificar direta ou indiretamente os indivíduos participantes do experimento, garantindo, assim, a sigiliosidade das informações coletadas. A partir deste momento, os participantes do experimento passaram a ser identificados, exclusivamente, por um número de identificação gerado aleatoriamente.

A tabela 9 exhibe alguns atributos que estão disponíveis e foram recuperados a partir dos *logs* do sistema de gerenciamento de aprendizagem (Moodle) da UFAL e que foram utilizados para a construção do modelo computacional proposto. A Tabela 10 exhibe as pontuações do cinco traços de personalidade, calculados a partir das respostas fornecidas pelos participante dos experimento.

Tabela 9 – Atributos utilizados pela pesquisa e que são recuperados a partir do sistema de gestão de aprendizagem da UFAL

<b>Atributo</b>	<b>Descrição</b>
genero	Gênero do aluno (0 - Masculino, 1 - Feminino)
num_periodos_cursados	Número de períodos cursados
mudou_foto_perfil	Informa se o aluno atualizou sua foto de perfil (0 - Não, 1 - Sim)
ajustou_autodescricao	Número de vezes que o aluno atualizou sua autodescrição na área de perfil
dias_interacao	Quantidade total de dias com acessos registrados no ambiente de aprendizagem
media_dias_interacao_periodo_letivo	Média de dias com acessos registrados no ambiente de aprendizagem por período letivo cursado
logins_realizados	Quantidade total de logins realizados no ambiente de aprendizagem
media_logins_realizados_periodo_letivo	Média de logins realizados no ambiente de aprendizagem
num_acessos_cursos	Quantidade total de acessos aos cursos realizados pelo aluno
media_acessos_cursos	Média de acessos aos cursos realizados pelo aluno por período letivo
num_acessos_materiais_cursos	Número de vezes que o aluno acessou os materiais do curso
media_acessos_materiais_cursos	Média de número de acessos realizados aluno aos materiais do curso
num_visualizacoes_atividade_foruns	Número de vezes que o usuário acessou atividades de fóruns

*Continua na próxima página*

Tabela 9 – *Continuação*

<b>Atributo</b>	<b>Descrição</b>
media_visualizacoes_atividade_foruns	Média de vezes que o usuário acessou atividades de fóruns por período letivo
num_visualizacoes_discussões_foruns	Número de vezes que o usuário acessou discussões de fóruns
media_visualizações_discussões_foruns	Média de vezes que o usuário acessou discussões de fóruns por período letivo
num_posts	Número de posts realizados nas atividades do tipo fórum
media_num_posts	Média de posts realizados nas atividades do tipo fórum por período letivo

Tabela 10 – Atributos utilizados pela pesquisa e que são obtidos a partir das respostas do questionário aplicado

<b>Atributo</b>	<b>Descrição</b>
pontuacao_abertura	Pontuação do traço de personalidade Abertura
pontuacao_conscienciosidade	Pontuação do traço de personalidade Conscienciosidade
pontuacao_extroversao	Pontuação do traço de personalidade Extroversão
pontuacao_amabilidade	Pontuação do traço de personalidade Amabilidade
pontuacao_neuroticismo	Pontuação do traço de personalidade Neuroticismo

## 4.2 Avaliação de Algoritmos Regressão para a Classificação de Traços de Personalidade

Durante a pesquisa foi detectada a necessidade de avaliar e comparar diferentes técnicas e algoritmos de mineração de dados a fim de quais deles são os mais eficientes para lidar com o tipo de problema considerado nesta investigação. Para este estudo serão empregados algoritmos de predição, mais especificamente algoritmos de regressão já que os escores dos cinco traços de personalidade apresentam valores numéricos contínuos.

### 4.2.1 Questões de investigação e Hipóteses

Esta proposta de pesquisa apresenta as seguintes perguntas de investigação:

**P1:** É possível estabelecer uma correlação entre os traços de personalidade dos alunos, aferidos pelos questionários aplicados, com o padrão de comportamento demonstrados pelos estudantes durante o uso do ambiente de aprendizagem à distância?

**H1.1-0:** Não há correlação significativa entre o traço de personalidade **Abertura** e as interações do estudante no AVA.

**H1.1-1:** Há correlação significativa entre o traço de personalidade **Abertura** e as interações do estudante no AVA.

**H1.2-0:** Não há correlação significativa entre o traço de personalidade **Conscienciosidade** e as interações do estudante no AVA.

**H1.2-1:** Há correlação significativa entre o traço de personalidade **Conscienciosidade** e as interações do estudante no AVA.

**H1.3-0:** Não há correlação significativa entre o traço de personalidade **Extroversão** e as interações do estudante no AVA.

**H1.3-1:** Há correlação significativa entre o traço de personalidade **Extroversão** e as interações do estudante no AVA.

**H1.4-0:** Não há correlação significativa entre o traço de personalidade **Amabilidade** e as interações do estudante no AVA.

**H1.4-1:** Há correlação significativa entre o traço de personalidade **Amabilidade** e as interações do estudante no AVA.

**H1.5-0:** Não há correlação significativa entre o traço de personalidade **Neuroticismo** e as interações do estudante no AVA.

**H1.5-1:** Há correlação significativa entre o traço de personalidade **Neuroticismo** e as interações do estudante no AVA.

**P2:** Qual é o algoritmo/técnica de regressão mais indicado para a temática de identificação da personalidade de alunos de ensino a distância? Em outras palavras, qual dos algoritmos observados apresenta melhor precisão e acurácia?

**H2-0:** O Coeficiente de Correlação para os algoritmos A e B é igual.

**H2-1:** O Coeficiente de Correlação para os algoritmos A e B é diferente.

**H3-0:** O Erro Quadrático Médio para os algoritmos de regressão é igual.

**H3-1:** O Erro Quadrático Médio para os algoritmos de regressão é diferente.

**H4-0:** O Desvio Absoluto Médio para os algoritmos de regressão é igual.

**H4-1:** O Desvio Absoluto Médio para os algoritmos de regressão é diferente.

#### 4.2.2 Fatores e Variáveis de Respostas

As variáveis independentes (e também fatores) utilizados no experimento são:

**Algoritmo de regressão:** Algoritmo de regressão que tem como entrada os dados dos questionários, respondidos pelos alunos e pelos logs do moodle dos respectivos alunos;

**Traços de personalidade:** A pontuação de cada um dos cinco traços de personalidade é obtida a partir das previsões realizadas pelos diferentes algoritmos de regressão. As pontuações estão de acordo com as faixas de valores e as definições do instrumento de pesquisa adotado por esse experimento 4.1.2.

#### Interações com o Ambiente Virtual de Aprendizagem: .

As variáveis de resposta são o Coeficiente de Correlação, o Desvio Absoluto Médio e o Erro Quadrático Médio; calculados ao final da execução de cada um dos algoritmos de regressão avaliados no experimento. Essas variáveis são descritas com mais detalhes na seção 4.2.6. À exceção da variável Algoritmo, todas as variáveis da pesquisa são numéricas.

#### 4.2.3 Níveis dos fatores

No caso do fator Técnica/Algoritmo, a execução do experimento é repetida para cada implementação de algoritmo de regressão, que se busca avaliar. Do mesmo modo, o fator Traço de Personalidade é organizado em sub-níveis, cada um representando um dos cinco dos traços de personalidade. A tabela 11 ilustra a definição desses níveis.

Tabela 11 – Experimento - Definição dos níveis de fatores

Fator	Níveis
Técnica/Algoritmo	Regressão Linear
	IBL
	M5Rules
	M5P
	Rede Neural MultilayerPerceptron SVM
Traço de personalidade	Abertura (Openness)
	Conscienciosidade (Conscientiousness)
	Extroversão (Extraversion)
	Amabilidade (Agreeableness) Neuroticismo (Neuroticism)

#### 4.2.4 Definição formal das hipóteses

As hipóteses desse experimento são definidas formalmente, conforme a tabela 12.

Tabela 12 – Experimento - Definição formal das hipóteses

Hipótese	Hipótese Nula	Hipótese Alternativa
H1.1	$H1.1_0 : \rho(A1) < 0.3$	$H1.1_1 : \rho(A1) \geq 0.3$
H1.2	$H1.2_0 : \rho(A2) < 0.3$	$H1.2_1 : \rho(A2) \geq 0.3$
H1.3	$H1.3_0 : \rho(A3) < 0.3$	$H1.3_1 : \rho(A3) \geq 0.3$
H1.4	$H1.4_0 : \rho(A4) < 0.3$	$H1.4_1 : \rho(A4) \geq 0.3$
H1.5	$H1.5_0 : \rho(A5) < 0.3$	$H1.5_1 : \rho(A5) \geq 0.3$
H2	$H2_0 : \overline{\rho(A1)} = \overline{\rho(A2)}$	$H2_1 : \overline{\rho(A1)} \neq \overline{\rho(A2)}$
H3	$H3_0 : \overline{EQM(A1)} = \overline{EQM(A2)}$	$H3_1 : \overline{EQM(A1)} \neq \overline{EQM(A2)}$
H4	$H4_0 : \overline{DMA(A1)} = \overline{DMA(A2)}$	$H4_1 : \overline{DMA(A1)} \neq \overline{DMA(A2)}$

Onde  $\rho$  é uma função que retorna o valor da métrica do coeficiente de correlação,  $EQM$  é uma função que calcula a métrica Erro Quadrático Médio e  $DAM$  é um acrônimo para Desvio Absoluto Médio, que é uma outra medida de precisão usada para quantificar a eficiência de algoritmos de regressão. Essas medidas de precisão são detalhadas na seção 4.2.6.

Já os termos **A1** e **A2** se referem aos algoritmos de regressão a serem comparados. Estes algoritmos - e suas entradas - estão definidos na tabela 11.

#### 4.2.5 Design do Experimento

Em um primeiro instante, esses dados foram submetidos a uma etapa de processamento por parte dos algoritmos de classificação para estabelecer modelos que buscam prever as pontuações dos traços de personalidade (Seção 2.1.5.1) obtidas pelos participantes do experimento.

Com o propósito de verificar o comportamento dos algoritmos sobre cada combinação possível dos dados de entrada (Tabela 9) e analisar as interações entre os fatores (Tabela 11), o design utilizado será o Design Fatorial Completo com 10 replicações. Ou seja, no total 30 ensaios serão realizados, cada um 10 vezes, totalizando 300 execuções. A Tabela 13 descreve sucintamente cada um dos ensaios executados neste experimento.

Em cada ensaio de execução do algoritmo de classificação, os dados de entrada são particionados através da técnica da validação cruzada (do inglês, *cross-validation*). De acordo com esse método, os dados de entrada são divididos em  $k$  partições mutuamente exclusivas (chamadas de *folds*), com tamanho aproximado de  $n \div k$ . As primeiras  $(k - 1)$  *folds* são usadas para treinar o algoritmo de classificação (para estabelecer o modelo) e o *fold* restante serve para avaliar o modelo gerado.

#### 4.2.6 Medidas de precisão

As variáveis de precisão permitem que o desempenho de um algoritmo de regressão seja avaliado e comparado com o de outros algoritmos. Nesta pesquisa, o objetivo é avaliar a

Tabela 13 – Definição das execuções dos ensaios do experimento

<b>Ensaio</b>	<b>Algoritmo</b>	<b>Traço de Personalidade</b>
1	Regressão Linear	Abertura
2	IBL	
3	M5Rules	
4	M5P	
5	Rede Neural MultilayerPerceptron	
6	SVM	
7	Regressão Linear	Conscienciosidade
8	IBL	
9	M5Rules	
10	M5P	
11	Rede Neural MultilayerPerceptron	
12	SVM	
13	Regressão Linear	Extroversão
14	IBL	
15	M5Rules	
16	M5P	
17	Rede Neural MultilayerPerceptron	
18	SVM	
19	Regressão Linear	Amabilidade
20	IBL	
21	M5Rules	
22	M5P	
23	Rede Neural MultilayerPerceptron	
24	SVM	
25	Regressão Linear	Neuroticismo
26	IBL	
27	M5Rules	
28	M5P	
29	Rede Neural Rede Neural MultilayerPerceptron	
30	SVM	

capacidade dos algoritmos em preverem uma variável numérica contínua, em particular, busca-se avaliar a capacidade de previsão das pontuações dos cinco traços de personalidade (Tabela 10). No caso específico de variáveis contínuas, as métricas de precisão são, em geral, obtidas comparando-se as previsões dos modelos com os valores reais das variáveis-alvo, e na sequência, estabelecendo a diferença média entre os valores previstos e os valores reais.

Para viabilizar a avaliação e a comparação dos algoritmos de regressão analisados neste experimento, três medidas de precisão foram escolhidas e são aplicadas na análise dos resultados dos ensaios de execução planejados para esta pesquisa (Tabela 13). Neste experimento, as três medidas de precisão adotadas são: Coeficiente de Correlação de Pearson (Correl), Erro

Quadrático Médio (EQM) e Desvio Absoluto Médio (DAM).

O coeficiente de correlação, mede a correlação estatística entre os valores previstos pelos algoritmos de regressão e os valores reais encontrados no conjunto de dados. O coeficiente de correlação varia de 1 para resultados perfeitamente correlacionados, a 0 quando não há correlação e  $-1$  quando os resultados estão perfeitamente correlacionados negativamente. A equação abaixo, descreve o cálculo da métrica de precisão.

$$\rho_{e_j}(A) = \frac{Cov(x, y)}{\sigma_x \sigma_y} \quad (4.6)$$

Onde,

$$Cov(x, y) = \frac{\sum_i^n (x_i - \bar{x})(y_i - \bar{y})}{n - 1}, \sigma_x = \frac{\sum_i^n (x_i - \bar{x})^2}{n - 1}, \sigma_y = \frac{\sum_i^n (y_i - \bar{y})^2}{n - 1} \quad (4.7)$$

Nas equações 4.6 e 4.7, os termos  $n$ ,  $x_i$  e  $y_i$ , representam, respectivamente, a quantidade de observações registradas no conjunto de dados de entrada, o valor da  $i$ -ésima observação real do conjunto de dados e o valor da  $i$ -ésima previsão realizada pelo algoritmo de regressão. O termo  $e_j$  representa a  $j$ -ésima execução do ensaio de execução do algoritmo analisado, este último simbolizado pelo componente A da equação. Cada ensaio deste experimento será executado dez vezes, então a variável  $\rho$  será calculada dez vezes para cada par de algoritmo de regressão e traço de personaliade listado na Tabela 13.

Para problemas de regressão, a medida Erro Quadrático Médio (EQM) é uma das métricas de precisão mais utilizadas, sendo empregada em uma ampla gama de problemas das mais diversas áreas de aplicação (JAMES et al., 2013). Essa métrica corresponde à média das diferenças (elevadas ao quadrado) encontradas entre os valores reais observados no conjunto de dados e os valores preditos pela função de regressão. A equação abaixo, resume a composição da medida.

$$EQM_{e_j}(A) = \frac{1}{n} \sum_n^{i=1} (y_i - \hat{f}(x_i))^2 \quad (4.8)$$

Onde  $n$  representa a quantidade de instâncias (ou observações) que compõem o conjunto de dados de entrada,  $y_i$  é um valor real encontrado na  $i$ -ésima posição do conjunto de dados,  $\hat{f}(x_i)$  representa o valor previsto pela função/algoritmo de previsão para a  $i$ -ésima observação. O termo  $e_j$  representa a  $j$ -ésima execução do ensaio de execução do algoritmo A analisado.

O valor da métrica  $EQM$  tenderá a ser pequeno quando os valores previstos estiverem muito próximos dos valores verdadeiros, e será grande se, para os conjuntos de dados, nos quais os valores previstos diferem substancialmente das observações verdadeiras.



Uma outra métrica de precisão utilizada no presente experimento é o Desvio Absoluto Médio (*DAM*). Para calcular a métrica DAM, tomamos a diferença absoluta entre os valores previstos e os valores reais do conjunto de dados. Então, tomamos a média dessas diferenças absolutas.

$$DAM_{e_j}(A) = \frac{1}{n} \sum_n^{i=1} |y_i - \hat{f}(x_i)| \quad (4.9)$$

A escolha da métrica de precisão depende da composição dos dados e da área da aplicação. Segundo Witten, Frank e Hall (2011), a *EQM* é uma boa métrica de avaliação para muitas aplicações, pois ela é de fácil manipulação matemática e tende a se comportar melhor, segundo os estatísticos, em um grande número de situações. No entanto, a métrica *EQM* é mais suscetível às diferenças provocadas por *outliers*, casos em que o erro da previsão é maior que os das demais observações. Por outro lado, a métrica *DAM* não apresenta essa desvantagem. Essa métrica é mais intuitiva e menos sensível a *outliers*.

Dados os valores de  $\rho(e_j)$ ,  $EQM(e_j)$  e  $DAM(e_i)$  para a  $j$ -ésima execução do algoritmo de regressão A. As médias das medidas  $\rho$ , DAM e EQM calculadas a respeito do algoritmo A são dadas pelas Equações 4.10, 4.11 e 4.12, em que  $j$  corresponde ao número de execuções do algoritmo A.

$$\bar{\rho}(A) = \frac{1}{n} \sum_n^{j=1} \rho(e_j) \quad (4.10)$$

$$\overline{EQM}(A) = \frac{1}{n} \sum_n^{j=1} EQM(e_j) \quad (4.11)$$

$$\overline{DAM}(A) = \frac{1}{n} \sum_n^{j=1} DAM(e_j) \quad (4.12)$$

E com os valores das médias em mãos, é possível calcular o desvio padrão ( $\sigma$ ) para as métricas  $\rho(e_j)$ ,  $EQM(e_j)$  e  $DAM(e_i)$  a partir das equações 4.13, 4.14 e 4.15, respectivamente.

$$\sigma_{\rho}(A) = \sqrt{\frac{1}{n-1} \sum_n^{j=1} (\rho_{e_j}(A) - \bar{\rho}(A))^2} \quad (4.13)$$

$$\sigma_{EQM}(A) = \sqrt{\frac{1}{n-1} \sum_n^{j=1} (EQM_{e_j}(A) - \overline{EQM}(A))^2} \quad (4.14)$$

$$\sigma_{DAM}(A) = \sqrt{\frac{1}{n-1} \sum_n^{j=1} (DAM_{e_j}(A) - \overline{DAM}(A))^2} \quad (4.15)$$

#### 4.2.7 Execução dos Algoritmos

Para efetuar as avaliações dos preditores apresentados anteriormente, essa pesquisa fez uso do pacote de software WEKA, desenvolvido como um projeto de código aberto, gerido pela Universidade de Waikato da Nova Zelândia (Machine Learning Group at the University of Waikato, 2019). O WEKA fornece uma extensa coleção de algoritmos de aprendizagem de máquina, prontos para serem usados; além de disponibilizar um conjunto de ferramentas projetadas para auxiliar o usuário durante todas as etapas do processo de mineração de dados, abrangendo, inclusive, as etapas de preparação dos dados de entrada, escolha e configuração dos algoritmos de aprendizagem, execução e avaliação da aprendizagem e visualização das saídas e resultados finais da aprendizagem de máquina (HALL et al., 2009). As figuras 1 e 2 exibem exemplos de tela do ambiente WEKA.



Figura 1 – Tela inicial do ambiente WEKA

A fim de verificar a acurácia das métricas definidas na seção 4.2.6, cada implementação dos algoritmos, analisados no estudo, foi executada 10 vezes para cada um dos cinco traços de personalidade. Sendo que em cada rodada de execução, há um elemento de aleatoriedade na maneira pela qual o conjunto de dados de entrada é particionado para as etapas de treinamento e de avaliação dos algoritmos de regressão. Essa aleatoriedade é proporcionada pela mudança do valor do parâmetro semente (*seed*), ajustada para cada nova execução.

Como em cada interação, o conjunto de dados de entrada é dividido em partições diferentes, as métricas  $\rho$  (Equação 4.6), EQM (Equação 4.8) e DAM (Equação 4.9) acabam por assumir valores ligeiramente distintos ao final de cada execução. No final do procedimento, os valores de cada interação são combinados para calcular as médias aritméticas das três medidas de precisão (Equações 4.10, 4.11 e 4.12).

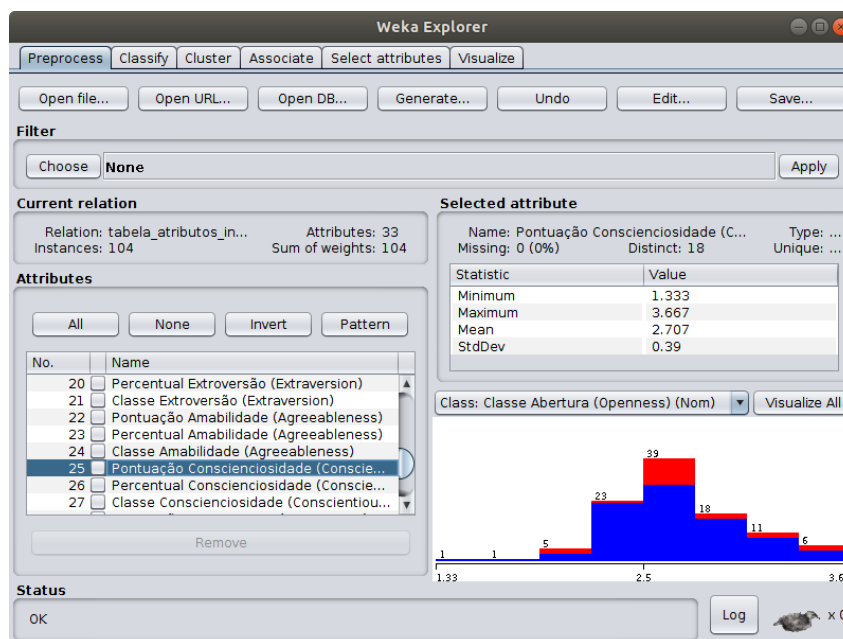


Figura 2 – Tela de seleção de atributos WEKA

#### 4.2.7.1 Regressão Linear

Quando o resultado, ou classe a ser predita, é numérica, e todos os atributos de entrada podem ser convertidos para números, a regressão linear geralmente é uma das primeiras opções a serem consideradas. Este é um método fundamental na estatística (WITTEN; FRANK; HALL, 2011).

A regressão linear é um método excelente e simples para previsão numérica e tem sido amplamente utilizado em várias aplicações por décadas. Porém, este algoritmo não é indicado para situações nas quais os conjuntos de dados analisados não estão distribuídos linearmente. Para esse tipo situação, o modelo criado a partir da regressão linear não correlacionará bem com os dados de entrada.

O Código 4.1 exibe a função criada para tentar prever a pontuação do traço de Neuroticismo. O algoritmo foi executado com sua configuração padrão.

#### Código 4.1 – Função de regressão linear gerada pelo WEKA

```

1 Linear Regression Model
2
3 Pontuacao Neuroticismo (Neuroticism) =
4
5     0.0202 * Quantidade de periodos letivos cursados +
6     -0.1899 * Mudou foto do perfil +
7     -0.0001 * Total de visualizacoes de cursos +
8     0.0005 * Media de visualizacoes de cursos por periodo cursado +
9     -0.0004 * Total de visualizacoes de recursos/materiais de cursos +
10    0.0034 * Media de visualizacoes de recursos/materiais de cursos por p
+

```

```

11      0.0004 * Total de visualizacoes de discussoes da atividade forum +
12      -0.0013 * Media de visualizacoes de discussoes da atividade forum po +
13      -0.0073 * Media de posts criados em discussoes da atividade forum por
      +
14      3.4098

```

---

#### 4.2.7.2 Instance-Based Learning (IBL)

A plataforma WEKA oferece o algoritmo IBK como alternativa de implementação para a técnica de Aprendizagem Baseada em Instâncias (da sigla em inglês IBL – *Instance-Based Learning*). O IBK pode ser categorizado como uma sub-classe do algoritmo de classificação K-vizinhos mais próximos (do inglês, *k-nearest neighbor* – *KNN*). Segundo Silva, Yamakami e Almeida (2012), algoritmos da classe KNN priorizam primordialmente manter uma consistência perfeita com o conjunto inicial de treinamento. Já os algoritmos IBL buscam otimizar a acurácia perante a novas instâncias do problema.

A ideia central por trás do funcionamento do algoritmos KNN e IBK é que em um espaço n-dimensional, um ponto P e os seus vizinhos mais próximos, pertencem a mesma classe. Ainda segundo Silva, Yamakami e Almeida (2012), assim como na técnica KNN, no método IBK existe uma função de similaridade que obtém um valor numérico obtido, geralmente, pelo cálculo da distância euclidiana. Então, os valores desta função de similaridade definem o quão próximos os dados da entrada estão próximos uns dos outros, e quais desses dados podem ser agrupados pertencem a mesma classe.

Um aspecto negativo desta técnica é que esta, ao final da etapa de treinamento, não disponibiliza um modelo de saída que explique como os padrões foram encontrados.

#### 4.2.7.3 M5Rules

O algoritmo M5Rules obtém regras de regressão a partir de árvores de decisão construídas por um outro algoritmo de regressão (algoritmo M5') (HOLMES; HALL; FRANK, 1999). O Código 4.2 exhibe um modelo de predição baseado em regras gerados pelo M5Rules. O algoritmo foi executado com sua configuração padrão.

---

#### Código 4.2 – Regras de regressão geradas pelo algoritmo M5Rules

---

```

1 M5 pruned model rules
2 (using smoothed linear models) :
3 Number of Rules : 2
4
5 Rule: 1
6 IF
7   Total de visualizacoes de recursos/materiais de cursos <= 629.5
8 THEN
9

```

```

10 Pontuacao Neuroticismo (Neuroticism) =
11 0.004 * Quantidade de periodos letivos cursados
12 - 0.0281 * Mudou foto do perfil
13 - 0.0009 * Total de dias de interacoes no Moodle
14 + 0.0031 * Media de dias de interacoes no Moodle por periodo cursado
15 - 0 * Total de visualizacoes de cursos
16 + 0.0001 * Media de visualizacoes de cursos por periodo cursado
17 + 0.0007 * Total de visualizacoes de recursos/materiais de cursos
18 + 0.0005 * Media de visualizacoes de recursos/materiais de cursos por p
19 + 0.0004 * Total de visualizacoes de discussoes da atividade forum
20 - 0.0021 * Media de visualizacoes de discussoes da atividade forum po
21 - 0.0001 * Total de posts criados em discussoes da atividade forum
22 + 3.2745 [82/95.464%]
23
24 Rule: 2
25
26 Pontuacao Neuroticismo (Neuroticism) =
27 -0.0606 * Quantidade de periodos letivos cursados
28 + 3.7682 [22/91.969%]

```

---

#### 4.2.7.4 M5P

O algoritmo M5P obtém as regressão a partir de árvores de decisão construídas por um outro algoritmo de regressão (algoritmo M5') (HOLMES; HALL; FRANK, 1999). O Código 4.3 exibe um modelo de predição baseado em regras gerados pelo M5Rules. O algoritmo foi executado com sua configuração padrão.

#### Código 4.3 – Regras de regressão geradas pelo algoritmo M5P

---

```

1 M5 pruned model tree:
2 (using smoothed linear models)
3
4 Total de visualizacoes de recursos/materiais de cursos <= 629.5 : LM1
   (82/95.464%)
5 Total de visualizacoes de recursos/materiais de cursos > 629.5 : LM2
   (22/61.958%)
6
7 LM num: 1
8 Pontuacao Neuroticismo (Neuroticism) =
9 0.004 * Quantidade de periodos letivos cursados (SIEWEB)
10 - 0.0281 * Mudou foto do perfil
11 - 0.0009 * Total de dias de interacoes no Moodle
12 + 0.0031 * Media de dias de interacoes no Moodle por periodo cursado
13 - 0 * Total de visualizacoes de cursos
14 + 0.0001 * Media de visualizacoes de cursos por periodo cursado
15 + 0.0007 * Total de visualizacoes de recursos/materiais de cursos
16 + 0.0005 * Media de visualizacoes de recursos/materiais de cursos por p

```

```

17 + 0.0004 * Total de visualizacoes de discussoes da atividade forum
18 - 0.0021 * Media de visualizacoes de discussoes da atividade forum po
19 - 0.0001 * Total de posts criados em discussoes da atividade forum
20 + 3.2745
21
22 LM num: 2
23 Pontuacao Neuroticismo (Neuroticism) =
24 -0.0254 * Quantidade de periodos letivos cursados (SIEWEB)
25 - 0.0737 * Mudou foto do perfil
26 - 0 * Total de visualizacoes de cursos
27 + 0.0002 * Media de visualizacoes de cursos por periodo cursado
28 - 0.0002 * Total de visualizacoes de recursos/materiais de cursos
29 + 0.0013 * Media de visualizacoes de recursos/materiais de cursos por p
30 + 0.0002 * Total de visualizacoes de discussoes da atividade forum
31 - 0.0007 * Media de visualizacoes de discussoes da atividade forum po
32 - 0.0004 * Total de posts criados em discussoes da atividade forum
33 + 3.6008
34
35 Number of Rules : 2

```

---

#### 4.2.7.5 Rede Neural Perceptron Multicamadas

Para treinar a rede neural artificial, foi utilizado o algoritmo MultilayerPerceptron do WEKA, que utiliza a função de backpropagation para classificar as instâncias. O algoritmo foi executado com a configuração padrão fornecida pela plataforma WEKA.

##### Código 4.4 – Parte do modelo de regressão gerado pelo algoritmo MultilayerPerceptron

```

1 === Classifier model (full training set) ===
2
3 Linear Node 0
4   Inputs      Weights
5   Threshold   0.013306816200475358
6   Node 1     -1.5387078574917632
7   Node 2     1.0720068796231061
8   Node 3     1.2428950277649313
9   Node 4     -1.758332025790462
10  Node 5     2.2357326712183245
11  Node 6     -1.26110170627393
12  Node 7     2.1090865617242063
13  Node 8     -1.6637863101095474
14  Node 9     -1.6886227594421503
15 Sigmoid Node 1
16  Inputs      Weights
17  Threshold   -0.41916756492005836
18  Attrib sexo=F    1.391400861839438
19  Attrib Quantidade de periodos letivos cursados    1.0733160105824608

```

20     Attrib Mudou foto do perfil     -1.7378502321972662  
21     Attrib Mudou descricao do perfil     2.396213789625132  
22     Attrib Total de dias de interacoes no Moodle     0.47555862953170946  
23     Attrib Media de dias de interacoes no Moodle por periodo cursado  
0.13543527327484411  
24     Attrib Total de logins realizados no Moodle     0.85322251314624  
25     Attrib Media de logins realizados no Moodle por periodo cursado  
0.30723987473298664  
26     Attrib Total de visualizacoes de cursos     2.139927199428465  
27     Attrib Media de visualizacoes de cursos por periodo cursado  
1.3526537795213085  
28     Attrib Total de visualizacoes de recursos/materiais de cursos  
1.2068543265916116  
29     Attrib Media de visualizacoes de recursos/materiais de cursos por p  
0.2569422689158814  
30     Attrib Total de visualizacoes da atividade forum     0.4135374515694422  
31     Attrib Media de visualizacoes da atividade forum por periodo curs  
0.6267266299578321  
32     Attrib Total de visualizacoes de discussoes da atividade forum  
0.2956698119601142  
33     Attrib Media de visualizacoes de discussoes da atividade forum po  
0.5387807286461034  
34     Attrib Total de posts criados em discussoes da atividade forum  
0.3704835572501595  
35     Attrib Media de posts criados em discussoes da atividade forum por  
-0.132951655422848  
36 Sigmoid Node 2  
37     Inputs     Weights  
38     Threshold     -0.552824452911355  
39     Attrib sexo=F     2.6617432540786474  
40     Attrib Quantidade de periodos letivos cursados     -1.584528349876831  
41     Attrib Mudou foto do perfil     -1.2031023070209874  
42     Attrib Mudou descricao do perfil     2.2865183160418945  
43     Attrib Total de dias de interacoes no Moodle     3.1678992567577446  
44     Attrib Media de dias de interacoes no Moodle por periodo cursado  
2.6687478514703784  
45     Attrib Total de logins realizados no Moodle     -0.8479949390759072  
46     Attrib Media de logins realizados no Moodle por periodo cursado  
-1.0911971228587887  
47     Attrib Total de visualizacoes de cursos     -1.4429864037625442  
48     Attrib Media de visualizacoes de cursos por periodo cursado  
-0.36506177684723734  
49     Attrib Total de visualizacoes de recursos/materiais de cursos  
-0.7713529203671445  
50     Attrib Media de visualizacoes de recursos/materiais de cursos por p  
0.5573688969057783  
51     Attrib Total de visualizacoes da atividade forum

```

-0.48846375296809846
52  Attrib Media de visualizacoes da atividade forum por periodo curs
0.24188601849565505
53  Attrib Total de visualizacoes de discussoes da atividade forum
1.714366872178924
54  Attrib Media de visualizacoes de discussoes da atividade forum po
0.6418923485748758
55  Attrib Total de posts criados em discussoes da atividade forum
-0.0928257036813367
56  Attrib Media de posts criados em discussoes da atividade forum por
0.08809661823398543

```

---

#### 4.2.7.6 Support Vector Machines (SVM)

O algoritmo SMOReg do WEKA fornece uma implementação da técnica de *Support Vector Machines*. O algoritmo foi executado com sua configuração padrão. O modelo de predição gerado pelo SMOReg pode ser visualizado no Código 4.5.

#### Código 4.5 – Regras de regressão geradas pelo algoritmo SMOReg

---

```

1 weights (not support vectors):
2  -      0.0042 * (normalized) sexo=F
3  +      0.0294 * (normalized) Quantidade de periodos letivos cursados
4  -      0.0379 * (normalized) Mudou foto do perfil
5  -      0.0637 * (normalized) Mudou descricao do perfil
6  -      0.1973 * (normalized) Total de dias de interacoes no Moodle
7  +      0.0451 * (normalized) Media de dias de interacoes no Moodle por
    periodo cursado
8  -      0.0387 * (normalized) Total de logins realizados no Moodle
9  -      0.0339 * (normalized) Media de logins realizados no Moodle por
    periodo cursado
10 -      0.2959 * (normalized) Total de visualizacoes de cursos
11 +      0.0659 * (normalized) Media de visualizacoes de cursos por periodo
    cursado
12 -      0.0438 * (normalized) Total de visualizacoes de recursos/materiais
    de cursos
13 +      0.1179 * (normalized) Media de visualizacoes de recursos/materiais
    de cursos por p
14 +      0.2892 * (normalized) Total de visualizacoes da atividade forum
15 -      0.1865 * (normalized) Media de visualizacoes da atividade forum
    por periodo curs
16 +      0.2504 * (normalized) Total de visualizacoes de discussoes da
    atividade forum
17 -      0.244  * (normalized) Media de visualizacoes de discussoes da
    atividade forum po
18 +      0.0615 * (normalized) Total de posts criados em discussoes da
    atividade forum

```



19 + 0.154 \* (normalized) Media de posts criados em discussoes da  
atividade forum por  
20 + 0.595

---



## 5 RESULTADOS

Para realizar a análise comparativa dos algoritmos de regressão, este trabalho segue a metodologia descrita por Dosualdo e Rezende (2003), que combina as equações das médias e desvios padrão de uma das métricas de precisão (Seção 4.2.6) na Equação 5.1, apresentada a seguir. Neste trabalho a métrica DAM foi a escolhida para o cálculo dos componentes média e desvio padrão, como pode ser observado nas Equações 5.2 e 5.3.

Os termos  $A_s$  e  $A_p$  fazem referência aos algoritmos que estão sendo comparados,  $A_s$  representa o algoritmo de regressão padrão S e o termo  $A_p$  corresponde ao algoritmo proposto P.

$$ad(A_s - A_p) = \frac{media(A_s - A_p)}{dp(A_s - A_p)} \quad (5.1)$$

Onde,

$$media(A_s - A_p) = \overline{DAM}(A_s) - \overline{DAM}(A_p) \quad (5.2)$$

$$dp(A_s - A_p) = \sigma_{DAM}(A_s) - \sigma_{DAM}(A_p) \quad (5.3)$$

Ainda de acordo com Dosualdo e Rezende (2003), o valor resultante de  $ad(A_s - A_p)$  deve ser interpretado da seguinte maneira: se  $ad(A_s - A_p)$  culminar em um valor positivo ( $> 0$ ), então o desempenho do algoritmo  $A_p$  supera a performance do algoritmo  $A_s$ , ao passo que se  $ad(A_s - A_p) \leq 0$ , então o algoritmo de regressão  $A_s$  é mais eficiente que o algoritmo  $A_p$ .

Em paralelo, a fim de confirmar se há diferenças significativas nos desempenhos dos algoritmos, foram realizados testes estatísticos do tipo *t-student* (DIEZ; BARR; CETINKAYA-RUNDEL, 2015; MONTGOMERY, 2012) nas médias obtidas para as métricas correlação ( $\rho$ ), Desvio Absoluto Médio (DAM) e Erro Quadrático Médio (EQM).

Os dados da amostra analisada atendem ao requisito da normalidade amostral. Os resultados dos testes de normalidade Kolmogorov-Smirnov, listados nas Tabelas 14, 16, 18, 20 e 22, indicam que as observações das medidas  $\rho$ , DAM e EQM seguem o tipo de distribuição normal. As tabelas 15, 17, 19, 21 e 23 apresentam os resultados dos testes de hipóteses dos resultados das execuções dos algoritmos para os traços de personalidade de Abertura, .

### 5.1 Traço de Personalidade Abertura

A tabela 14 apresenta as médias das métricas de desempenho obtidas pelas implementações dos seis algoritmos de regressão, na tentativa de realizar a predição da pontuação do traço de personalidade abertura, a partir do conjunto de dados de entrada e de acordo com a proposta dos ensaios previstos para este experimento.

Tabela 14 – Resultados das métricas de precisão para os algoritmos avaliados pelo experimento para o traço de personalidade Abertura

Algoritmo	Coeficiente de Correlação ( $\rho$ )		DAM		EQM	
	Média	Teste KS	Média	Teste KS	Média	Teste KS
Regressão Linear	-0,1333	0,6519	0,3434	0,9762	0,1994	0,9255
IBL	0,0280	0,998	0,3987	0,8032	0,2624	0,8596
M5Rules	-0,0190	0,3118	0,3211	0,907	0,1745	0,6553
M5P	-0,0319	0,7108	0,3150	0,9882	0,1696	0,9018
MultilayerPerceptron	-0,0060	0,5771	0,5052	0,9466	0,4768	0,4811
SVM	0,0190	0,9909	0,3332	0,9826	0,1732	0,9536

A partir dos valores das métricas  $\rho$ , DAM e EQM, exibidas na Tabela 14, é possível perceber que nenhum dos algoritmos obteve desempenho satisfatório na tarefa de prever a pontuação do traço de personalidade Abertura. Dentre as opções sondadas, as previsões fornecidas pelo algoritmo IBL apresentam maior correlação com as pontuações reais, extraídas a partir das respostas do questionário de personalidade. Em relação às medidas DAM e EQM, o desempenho dos preditores também não se mostrou convincente. A média dos erros se aproxima do desvio padrão da pontuação Abertura na amostra analisada, de acordo com a estatística descritiva da Tabela 8.

A tabela 15 exhibe os resultados das comparações dos pares de preditores. Os testes estatísticos mostram que, em relação às métricas DAM e EQM, os algoritmos de regressão apresentam desempenhos diferentes; com exceção dos algoritmos M5Rules, M5P e MultilayerPerceptron. Quando a métrica coeficiente de correlação é considerada, o algoritmo IBL apresenta, com significância estatística, desempenho diferente em relação aos demais algoritmos.

Tabela 15 – Testes de Hipóteses dos resultados dos algoritmos para predição da pontuação de Traço de Personalidade Abertura

Algoritmo $A_S$	Algoritmo $A_P$	Medida $ad(A_S - A_P)$	Melhor Algoritmo	Teste Estatístico* (p-valor)		
				Coef. ( $\rho$ )	DAM	EQM
Regressão Linear	IBL	6,4498	Regressão Linear	$1,634^{-7}$	$4,353^{-8}$	$2,959^{-9}$
Regressão Linear	M5Rules	-1,5158	M5Rules	0,1051	0,0038	0,0004
Regressão Linear	M5P	-2,4940	M5P	0,1706	0,0001	$9,533^{-6}$
Regressão Linear	MultilayerPerceptron	4,7609	Regressão Linear	0,0767	$2,26^{-7}$	$4,857^{-6}$
Regressão Linear	SVM	-2,3514	SVM	0,0098	0,0028	$1,019^{-5}$
IBL	M5Rules	-12,6808	M5Rules	$1,437^{-7}$	$7,222^{-9}$	$7,516^{-11}$
IBL	M5P	-29,7929	M5P	$1,08^{-7}$	$1,011^{-10}$	$4,128^{-12}$
IBL	MultilayerPerceptron	4,1913	IBL	$6,368^{-7}$	$3,842^{-6}$	$3,629^{-5}$
IBL	SVM	15,3377	IBL	$7,67^{-11}$	$2,636^{-10}$	$5,938^{-12}$
M5Rules	M5P	1,8500	M5Rules	0,7603	0,4239	0,4347
M5Rules	MultilayerPerceptron	9,5370	M5Rules	0,7764	$3,321^{-9}$	$1,875^{-6}$
M5Rules	SVM	-1,1692	SVM	0,2663	0,0773	0,8194
M5P	MultilayerPerceptron	8,4126	M5P	0,5709	$5,376^{-9}$	$1,813^{-6}$
M5P	SVM	-2,5809	SVM	0,1431	0,0043	0,4963
MultilayerPerceptron	SVM	5,7929	MultilayerPerceptron	0,5052	$7,62^{-8}$	$2,1^{-6}$

\* Considerando um intervalo de confiança de 95%

## 5.2 Traço de Personalidade Conscienciosidade

Ao analisar os dados da tabela 16, observa-se que nenhum dos algoritmos alcançou resultados razoáveis ao tentar prever as pontuações do traço de Conscienciosidade. Tendo em vista a métrica coeficiente de correlação, o algoritmo de predição SVM obteve um índice de correlação maior que os demais preditores; mesmo assim, a correlação das previsões fornecidas se mostrou muito fraca para ser considerada relevante.

Tabela 16 – Resultados das métricas de precisão para os algoritmos avaliados pelo experimento para o traço de personalidade Conscienciosidade

Algoritmo	Coeficiente de Correlação ( $\rho$ )		DAM		EQM	
	Média	Teste KS	Média	Teste KS	Média	Teste KS
Regressão Linear	0,1511	0,7386	0,3255	0,8849	0,1863	0,8566
IBL	0,2460	0,7971	0,3742	0,9682	0,2314	0,5826
M5Rules	0,1989	0,7666	0,3137	0,8865	0,1669	0,4575
M5P	0,2069	0,46	0,3103	0,9331	0,1648	0,6502
MultilayerPerceptron	0,1161	0,9976	0,5356	0,9867	0,5055	0,7121
SVM	0,2980	0,9528	0,2907	0,8769	0,1489	0,9701

A tabela 17 exibe os resultados das conferências dos resultados alcançados pelos preditores para a pontuação do traço de personalidade Conscienciosidade. Os testes estatísticos mostram que, em relação às métricas DAM e EQM, os algoritmos de regressão apresentam desempenhos diferentes; com exceção dos algoritmos Regressão Linear, M5Rules e M5P que mostraram desempenhos semelhantes. Em relação à métrica coeficiente de correlação, o algoritmo SVM apresentou desempenho diferente em relação aos demais algoritmos.

Tabela 17 – Testes de Hipóteses dos resultados dos algoritmos para predição da pontuação de Traço de Personalidade Conscienciosidade

Algoritmo $A_S$	Algoritmo $A_P$	Medida $ad(A_S - A_P)$	Melhor Algoritmo	Teste Estatístico* (p-valor)		
				Coef. ( $\rho$ )	DAM	EQM
Regressão Linear	IBL	42,7238	Regressão Linear	$9,6355^{-5}$	$1,1312^{-8}$	$6,7410^{-8}$
Regressão Linear	M5Rules	-1,8274	M5Rules	0,0813	0,0799	0,0038
Regressão Linear	M5P	-2,1907	M5P	0,0392	0,0316	0,0034
Regressão Linear	MultilayerPerceptron	6,9658	Regressão Linear	0,1840	$1,3608^{-7}$	$1,2807^{-5}$
Regressão Linear	SVM	13,8289	Regressão Linear	$2,0395^{-7}$	$1,9612^{-7}$	$5,0678^{-8}$
IBL	M5Rules	-11,4295	M5Rules	0,0806	$7,0168^{-8}$	$6,1455^{-9}$
IBL	M5P	-11,0434	M5P	0,1290	$5,0220^{-8}$	$2,7091^{-8}$
IBL	MultilayerPerceptron	5,5613	IBL	$9,3016^{-5}$	$1,0646^{-6}$	$4,2744^{-5}$
IBL	SVM	22,8383	IBL	0,0062	$2,5795^{-12}$	$1,4751^{-11}$
M5Rules	M5P	-6,9138	M5P	0,7897	0,6597	0,7652
M5Rules	MultilayerPerceptron	9,3476	M5Rules	0,0129	$2,3154^{-8}$	$7,2889^{-6}$
M5Rules	SVM	2,5764	M5Rules	0,0011	0,0018	0,0047
M5P	MultilayerPerceptron	9,69376393444565	M5P	0,0062	$1,7774^{-8}$	$6,7419^{-6}$
M5P	SVM	2,0792	M5P	0,0016	0,0064	0,0170
MultilayerPerceptron	SVM	7,4942	MultilayerPerceptron	$2,726^{-6}$	$5,6149^{-8}$	$5,2872^{-6}$

\* Considerando um intervalo de confiança de 95%

### 5.3 Traço de Personalidade Extroversão

A Tabela 18 mostra os valores das métricas atingidos pelos algoritmos de regressão. Nenhum dos algoritmos demonstrou desempenho relevante na predição das pontuações do traço de personalidade Extroversão. O algoritmo IBL conseguiu um coeficiente de correlação maior que outros preditores. Mas o valor do índice de correlação, identificado durante o processo de avaliação dos algoritmos, repercutiu em uma correlação muito baixa.

Tabela 18 – Resultados das métricas de precisão para os algoritmos avaliados pelo experimento para o traço de personalidade Extroversão

Algoritmo	Coeficiente de Correlação ( $\rho$ )		DAM		EQM	
	Média	Teste KS	Média	Teste KS	Média	Teste KS
Regressão Linear	0,0490	0,6311	0,3456	0,887	0,1848	0,9123
IBL	0,2512	0,5725	0,3680	0,6671	0,2427	0,0556
M5Rules	0,0181	0,3388	0,3394	0,8768	0,1833	0,5192
M5P	0,0370	0,5545	0,3371	0,9901	0,1804	0,7088
MultilayerPerceptron	0,0845	0,582	0,5659	0,9896	0,6133	0,388
SVM	0,1732	0,02286*	0,3448	0,9912	0,1848	0,9527

A Tabela 19 exibe os resultados das conferências dos resultados alcançados pelos preditores para a pontuação do traço de personalidade Extroversão. Os testes estatísticos mostram que, em relação às métricas DAM e EQM, os algoritmos de regressão apresentam desempenhos diferentes; com exceção dos pares de preditores Regressão Linear – M5Rules, Regressão Linear – SVM, M5Rules – M5P e M5Rules – SVM, cujas médias das métricas de precisão revelaram-se muito similares. Em relação à métrica coeficiente de correlação, o algoritmo IBL apresentou desempenho dispare em relação dos outros preditores.

Tabela 19 – Testes de Hipóteses dos resultados dos algoritmos para predição da pontuação de Traço de Personalidade Extroversão

Algoritmo $A_S$	Algoritmo $A_P$	Medida $ad(A_S - A_P)$	Melhor Algoritmo	Teste Estatístico* (p-valor)		
				Coef. ( $\rho$ )	DAM	EQM
Regressão Linear	IBL	18,4201	Regressão Linear	$2,4433^{-8}$	$2,7415^{-6}$	0,0187
Regressão Linear	M5Rules	-3,3001	M5Rules	0,1925	0,0912	0,7247
Regressão Linear	M5P	-4,6923	M5P	0,5468	0,0244	0,2848
Regressão Linear	MultilayerPerceptron	7,1448	Regressão Linear	0,1883	$5,6775^{-7}$	$3,7390^{-5}$
Regressão Linear	SVM	0,4550	Regressão Linear	0,1289	0,7609	0,9925
IBL	M5Rules	-43,02940	M5Rules	$3,2487^{-8}$	$4,4146^{-7}$	0,0164
IBL	M5P	-51,6254	M5P	$6,0525^{-10}$	$1,3729^{-7}$	0,0130
IBL	MultilayerPerceptron	6,6815	IBL	$1,1946^{-5}$	$1,3534^{-6}$	$7,2511^{-5}$
IBL	SVM	7,6953	IBL	0,3166	$1,1094^{-6}$	0,0185
M5Rules	M5P	34,9979	M5Rules	0,3885	0,5559	0,4637
M5Rules	MultilayerPerceptron	7,8238	M5Rules	0,0251	$3,9408^{-7}$	$3,6448^{-5}$
M5Rules	SVM	-1,4631	SVM	0,0665	0,1086	0,6799
M5P	MultilayerPerceptron	7,8857	M5P	0,0714	$3,6171^{-7}$	$3,4805^{-5}$
M5P	SVM	-2,1283	SVM	0,0984	0,0271	0,1871
MultilayerPerceptron	SVM	6,7753	MultilayerPerceptron	0,2708	$5,9990^{-7}$	$3,7809^{-5}$

\* Considerando um intervalo de confiança de 95%

#### 5.4 Traço de Personalidade Amabilidade

A Tabela 20 mostra os valores das métricas obtidos pelos algoritmos de regressão analisados. Nenhum dos algoritmos demonstrou desempenho relevante na predição das pontuações do traço de personalidade Extroversão. O algoritmo IBL conseguiu um coeficiente de correlação maior que outros preditores, mas o valor alcançado é muito baixo, o que inviabiliza o uso do algoritmo em cenários realistas.

Tabela 20 – Resultados das métricas de precisão para os algoritmos avaliados pelo experimento para o traço de personalidade Amabilidade

Algoritmo	Coeficiente de Correlação ( $\rho$ )		DAM		EQM	
	Média	Teste KS	Média	Teste KS	Média	Teste KS
Regressão Linear	0,0088	0,8257	0,3392	0,728	0,1904	0,9582
IBL	0,1615	0,7557	0,3569	0,9986	0,2135	0,9959
M5Rules	-0,1341	0,05764	0,3225	0,9153	0,1728	0,9764
M5P	-0,0511	0,933	0,3205	0,5902	0,1702	0,7462
MultilayerPerceptron	0,0241	0,6814	0,5199	0,6738	0,5312	0,4712
SVM	-0,0936	0.04008	0,3280	0,5109	0,1803	0,9145

A Tabela 21 exibe os resultados das conferências dos resultados alcançados pelos preditores para a pontuação do traço de personalidade Amabilidade. Os testes estatísticos mostram que, em relação às métricas DAM e EQM, os algoritmos de regressão apresentam desempenhos diferentes; com exceção dos pares de preditores Regressão Linear – M5Rules, Regressão Linear – SVM, M5Rules – M5P e M5Rules – SVM, para a métrica DAM; e o par M5Rules – M5P para a medida de precisão EQM.

Tabela 21 – Testes de Hipóteses dos resultados dos algoritmos para predição da pontuação de Traço de Personalidade Amabilidade

Algoritmo $A_S$	Algoritmo $A_P$	Medida $ad(A_S - A_P)$	Melhor Algoritmo	Teste Estatístico* (p-valor)		
				Coef. ( $\rho$ )	DAM	EQM
Regressão Linear	IBL	2,0679	Regressão Linear	0,0001	0,0005	0,0004
Regressão Linear	M5Rules	-1,1347	M5Rules	0,1364	0,0005	0,0002
Regressão Linear	M5P	-1,6436	M5P	0,0834	$9,5788^{-5}$	$9,0298^{-6}$
Regressão Linear	MultilayerPerceptron	5,3147	Regressão Linear	0,0341	$8,7803^{-9}$	$1,2558^{-7}$
Regressão Linear	SVM	-2,1112	SVM	0,2061	0,0050	0,0135
IBL	M5Rules	-5,6235	M5Rules	0,0067	$1,1720^{-6}$	$1,086^{-6}$
IBL	M5P	-12,9723	M5P	$6,425^{-8}$	$3,5076^{-7}$	$7,0084^{-7}$
IBL	MultilayerPerceptron	6,4097	IBL	0,8021	$6,7719^{-9}$	$1,6118^{-7}$
IBL	SVM	8,8658	IBL	0,0060	$5,9132^{-6}$	$1,2113^{-5}$
M5Rules	M5P	0,6166	M5Rules	0,3607	0,6512	0,4855
M5Rules	MultilayerPerceptron	10,2223	M5Rules	0,0096	$1,7246^{-9}$	$7,9112^{-8}$
M5Rules	SVM	-0,5820	SVM	0,7189	0,2224	0,0770
M5P	MultilayerPerceptron	8,8172	M5P	0,0068	$2,0274^{-9}$	$8,1749^{-8}$
M5P	SVM	-1,2349	SVM	0,5774	0,0846	0,0105
MultilayerPerceptron	SVM	6,6891	MultilayerPerceptron	0,0118	$3,2098^{-9}$	$9,5887^{-8}$

\* Considerando um intervalo de confiança de 95%

Em relação à métrica coeficiente de correlação, foram identificadas diferenças significativas na performance entre alguns dos algoritmos de regressão estudados.

### 5.5 Traço de Personalidade Neuroticismo

A Tabela 22 mostra os valores das métricas obtidos pelos algoritmos de regressão analisados. Nenhum dos algoritmos demonstrou desempenho relevante na predição das pontuações do traço de personalidade Extroversão. O algoritmo IBL conseguiu um coeficiente de correlação maior que outros preditores, mas o valor alcançado é muito baixo, o que inviabiliza o uso do algoritmo em cenários realistas.

Tabela 22 – Resultados das métricas de precisão para os algoritmos avaliados pelo experimento para o traço de personalidade Neuroticismo

Algoritmo	Coeficiente de Correlação ( $\rho$ )		DAM		EQM	
	Média	Teste KS	Média	Teste KS	Média	Teste KS
Regressão Linear	-0,0380	0,2182	0,3415	0,002217*	0,1961	0,8651
IBL	0,1457	0,9841	0,4074	0,9534	0,2461	0,0688
M5Rules	-0,0920	0,3008	0,3253	0,9906	0,1810	0,9924
M5P	-0,0311	0,9782	0,3225	0,8659	0,1778	0,626
MultilayerPerceptron	0,1789	0,2782	0,5621	0,461	0,5704	0,5488
SVM	0,0477	0,9723	0,3300	0,8573	0,1819	0,677

A Tabela 23 exhibe os resultados das conferências dos resultados alcançados pelos preditores para a pontuação do traço de personalidade Neuroticismo. Os testes estatísticos mostram que, em relação às métricas DAM e EQM, os algoritmos de regressão apresentam desempenhos diferentes; com exceção dos pares de preditores M5Rules – M5P para as métricas DAM e EQM; e, M5Rules – SVM e M5P – SVM, apenas para a métrica EQM.



Tabela 23 – Testes de Hipóteses dos resultados dos algoritmos para predição da pontuação de Traço de Personalidade Neuroticismo

Algoritmo $A_S$	Algoritmo $A_P$	Medida $ad(A_S - A_P)$	Melhor Algoritmo	Teste Estatístico* (p-valor)		
				Coef. ( $\rho$ )	DAM	EQM
Regressão Linear	IBL	68,5152	Regressão Linear	0,0691	0,3410	0,0006
Regressão Linear	M5Rules	11,6431	Regressão Linear	0,2899	0,3409	0,0006
Regressão Linear	M5P	9,0130	Regressão Linear	0,5885	0,3409	$5,2170^{-5}$
Regressão Linear	MultilayerPerceptron	7,3976	Regressão Linear	0,0846	0,3412	$1,3747^{-7}$
Regressão Linear	SVM	3,5971	Regressão Linear	0,5693	0,3409	0,0005
IBL	M5Rules	34,9606	IBL	0,0040	$2,838^{-14}$	$9,1696^{-5}$
IBL	M5P	27,7215	IBL	$2,8499^{-6}$	$3,957^{-14}$	$6,8816^{-5}$
IBL	MultilayerPerceptron	5,3589	IBL	0,6270	$1,796^{-7}$	$9,5933^{-8}$
IBL	SVM	18,6116	IBL	$1,5189^{-5}$	$9,976^{-13}$	0,0001
M5Rules	M5P	3,9086	M5Rules	0,3694	0,3132	0,2958
M5Rules	MultilayerPerceptron	7,5861	M5Rules	0,0078	$5,055^{-9}$	$1,0225^{-7}$
M5Rules	SVM	-2,5463	SVM	0,0533	0,0795	0,7244
M5P	MultilayerPerceptron	7,5039	M5P	0,0113	$5,160^{-9}$	$1,0071^{-7}$
M5P	SVM	-6,7328	SVM	0,0029	0,0048	0,0831
MultilayerPerceptron	SVM	7,0294	MultilayerPerceptron	0,0800	$8,145^{-9}$	$1,1446^{-7}$

\* Considerando um intervalo de confiança de 95%



## 6 CONCLUSÃO

Neste capítulo faremos um breve resumo do trabalho e apresentaremos os nossos resultados com base no experimento realizado. Abordaremos, ainda, as principais limitações enfrentadas, bem como as principais ameaças à validade dos nossos resultados. Concluiremos, apresentando nossas principais contribuições.

Neste trabalho, argumentamos sobre a importância do conhecimento sobre os traços de personalidade dos estudantes com o objetivo de compreender, explicar e prever padrões comportamentais que os indivíduos externalizam quando vivenciam situações como: tomar decisões; estabelecer relações interpessoais; lidar com questões financeiras; gerir carreira, trabalho e estudo; dentre outras situações.

No contexto da educação, o conhecimento do traços de personalidade dos estudantes se justifica devido ao fato de estudos mostrarem a existência de correlação entre a personalidade dos estudantes e as pontuações obtidas em avaliações da aprendizagem. Dessa forma, percebemos ser possível, ao conhecer antecipadamente o tipo de personalidade de um aluno, estimar sucesso em atividades e avaliações, probabilidade de conclusão do curso, além de várias outras informações que podem ser geradas com o objetivo de ajudar o estudante, professores e gestores de cursos online. Sendo assim, o objetivo desta pesquisa foi o de identificar os traços de personalidade de cada estudante, utilizando, apenas, os dados oriundos das interações desses estudantes com o ambiente online de aprendizagem.

Para atingirmos esse objetivo, utilizamos a teoria *Big Five*, e o instrumento (questionário) *Big-5 Inventory* (Vide Anexo A) que, como resultado, informa a pontuação do indivíduo em cada um dos 5 traços de personalidade considerados: (1) Abertura à experiência; (2) Conscienciosidade; (3) Extroversão; (4) Amabilidade e; (5) Neuroticismo ou Instabilidade Emocional. O questionário e o termo de consentimento livre e esclarecido (TCLE) foi enviado e respondido por 113 estudantes de 9 cursos EAD. Após 90 dias, analisamos as respostas dos estudantes/participantes, mensurando cada traço de personalidade daqueles que responderam corretamente o questionário. Posteriormente, coletamos os dados educacionais dos participantes e contabilizamos as interações de cada tipo, para cada estudante. Por fim, realizamos uma análise de correlação entre os dados de interação e os traços de personalidade (conforme detectados pelo instrumento *Big-5 Inventory*), para investigar questionamentos do tipo: há correlação entre a maneira e a frequência do uso do ambiente de ensino com os traços de personalidade dos estudantes.

Como resultado, 108 estudantes preencheram corretamente o questionário. Entretanto, o resultado da análise de correlação mostrou não haver correlação significativa entre as interações analisadas (coletadas do ambiente online de aprendizagem **Moodle**) e os traços de personalidade (segundo a teoria *Big Five*). Sendo assim, concluímos que, seguindo a metodologia utilizada

nesta pesquisa, não é possível mensurar, automaticamente, os traços de personalidade com base nas interações realizadas pelos estudantes em um ambiente *Moodle*.

## 6.1 Limitações

Esta pesquisa sofreu com algumas limitações, por exemplo: calculamos que o tamanho ideal da amostra seria de 244 participantes, mas não conseguimos esse valor, mesmo após repetirmos por 4 vezes o envio do pedido de participação e de termos incluído outros cursos; O processo para a avaliação do projeto de pesquisa, pelo conselho de ética da Universidade Federal de Alagoas, ainda que compreendido e apoiado, demandou tempo que nos foi precioso; O curto tempo para realizar todas as atividades do mestrado (disciplinas, pesquisa, análise, escrita e apresentação), também, limitou a busca por outras formas de analisar os dados quando nos deparamos com os insucessos das análises realizadas;

Outro fator foi que o ambiente de aprendizagem utilizado (*Moodle*), não possuía uma quantidade e diversidade de dados necessários para a análise desejada. Finalmente, não tínhamos, para cada estudante analisado, nem quantidade, nem diversidade suficientes de interações que nos permitissem identificar, de forma inequívoca, seus traços de personalidade. O *Moodle* não oferece recursos que permitam ao aluno expressar a sua personalidade adequadamente. A organização do ambiente é inflexível e acaba por tratar todos alunos como iguais.

Algumas circunstâncias da pesquisa podem ameaçar a validade dos resultados. Por exemplo: utilizamos dados desbalanceados, agrupamos estudantes de diferentes cursos, que seguiam diferentes metodologias. Além de diferentes disciplinas em diferentes períodos do curso. Na nossa análise, e isso pode ter adicionado algum viés ao resultado.

## 6.2 Trabalhos Futuros

Dada a relevância dos traços de personalidade para a educação e tendo constatado que a arquitetura do *Moodle* é muito engessada; seria interessante, como um possível trabalho futuro, sugerir ajustes no *Moodle* ou desenvolver outras alternativas de AVA, que viabilizem ambientes de aprendizagem mais adaptáveis e que atentem para a importância dos traços de personalidade na aprendizagem. Atualmente os ambientes de aprendizagem não consideram o papel e a influência da personalidade dos indivíduos na educação.

Também como trabalho futuro, buscar combinar os registros das interações realizadas no *Moodle* com outras fontes de dados. Por exemplo, levar em conta os resultados das médias finais das disciplinas, registrados no sistema acadêmico da instituição e as interações realizadas em redes sociais como Facebook e Twitter.

Por fim, acreditamos que esta pesquisa, ainda que não tenha alcançado os objetivos pretendidos, contribuiu com os primeiros passos para o uso dos dados gerados por ambientes online de aprendizagem para a detecção automática dos traços de personalidade dos estudantes.

Outra contribuição é a nossa metodologia que pode ser adaptada, considerando as limitações e ameaças do nosso experimento.



## REFERÊNCIAS

- ALLPORT, G. W. *Personality: a psychological interpretation*. 1. ed. [S.l.]: Henry Holt And Company, 1937. Citado 3 vezes nas páginas 13, 18 e 19.
- ALLPORT, G. W.; ODBERT, H. Trait names: a psycho-lexical study. *Psychological Monographs*, v. 47, n. 211, 1936. Citado 2 vezes nas páginas 19 e 20.
- ANDRADE, J. M. de. *Evidências de Validade do Inventário dos Cinco Grandes Fatores de Personalidade para o Brasil*. Tese (Doutorado) — Universidade de Brasília, 2008. Citado 8 vezes nas páginas 13, 19, 20, 21, 23, 35, 37 e 38.
- ARIANI, D. W. Personality and learning motivation. *European Journal of Business and Management*, v. 5, n. 10, 2013. Citado na página 35.
- ASSUMPÇÃO, C. M. *Censo EAD.BR: relatório analítico da aprendizagem a distância no Brasil 2016*. [S.l.], 2017. Disponível em: <<http://www.abed.org.br>>. Citado na página 13.
- BAKER, R. et al. Better to be frustrated than bored: The incidence and persistence of affect during interactions with three different computer-based learning environments. *International Journal of human-computer studies*, v. 68, n. 4, p. 223–241, 2010. Citado na página 25.
- BARRICK, M. R.; MOUNT, M. K. The big five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, v. 44, p. 1–26, 03 1991. Citado na página 17.
- BENET, V.; JOHN, O. P. Los cinco grandes across cultures and ethnic groups: Multitrait multimethod analyses of the big five in spanish and english. *Journal of personality and social psychology*, v. 75, p. 729–50, 10 1998. Citado 2 vezes nas páginas 23 e 37.
- BIBBY, P.; FERGUSON, E. The ability to process emotional information predicts loss aversion. *Personality and Individual Differences*, v. 51, p. 263–266, 08 2011. Citado na página 17.
- BIDJERANO, T.; DAI, D. Y. The relationship between the big-five model of personality and self-regulated learning strategies. *Learning and Individual Differences*, v. 17, n. 1, p. 69–81, 2007. ISSN 1041-6080. Citado na página 24.
- BOGG, T.; ROBERTS, B. Conscientiousness and health-related behaviors: A meta-analysis of the leading behavioral contributors to mortality. *Psychological bulletin*, v. 130, p. 887–919, 11 2004. Citado na página 17.
- BOGG, T.; ROBERTS, B. The case for conscientiousness: Evidence and implications for a personality trait marker of health and longevity. *Annals of behavioral medicine : a publication of the Society of Behavioral Medicine*, v. 45, 12 2012. Citado na página 17.
- BRITO-COSTA, S. et al. Psychometric properties of portuguese version of big five inventory (bfi). *International Journal of Developmental and Educational Psychology. Revista INFAD de Psicología.*, v. 1, p. 83, 07 2016. Citado na página 23.
- BUSATO, V. et al. Intellectual ability, learning style, personality, achievement motivation and academic success of psychology students in higher education. *Personality and Individual Differences*, v. 29, n. 6, p. 1057 – 1068, 2000. ISSN 0191-8869. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0191886999002536>>. Citado na página 35.

CASPI, A. et al. The influence of personality on social participation in learning environments. *Learning and Individual Differences*, v. 16, n. 2, p. 129 – 144, 2006. Citado na página 24.

CATTELL, R. B. The description of personality: Basic traits resolved into clusters. *The journal of abnormal and social psychology*, American Psychological Association, v. 38, n. 4, p. 476, 1943. Citado na página 20.

CHEN, G. et al. On the impact of personality in massive open online learning. In: *Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization*. New York, NY, USA: ACM, 2016. (UMAP '16), p. 121–130. ISBN 978-1-4503-4368-8. Disponível em: <<http://doi.acm.org/10.1145/2930238.2930240>>. Citado 4 vezes nas páginas 9, 27, 29 e 30.

COCEA, M.; WEIBELZAHN, S. Disengagement detection in online learning: Validation studies and perspectives. *IEEE Transactions on Learning Technologies*, IEEE, v. 4, n. 2, p. 114–124, 2011. Citado na página 25.

CONARD, M. A. Aptitude is not enough: How personality and behavior predict academic performance. *Journal of Research in Personality*, v. 40, n. 3, p. 339 – 346, 2006. ISSN 0092-6566. Citado na página 24.

COSTA, P.; MCCRAE, R. R. Four ways five factors are basic. *Personality and Individual Differences*, v. 13, p. 653–665, 06 1992. Citado na página 21.

COSTA, P.; MCCRAE, R. R. *Revised NEO personality inventory (NEO PI-R) and NEO five-factor inventory (NEO-FFI) professional manual*. [S.l.]: Odessa, Fla. : Psychological Assessment Resources, 1992. Citado 2 vezes nas páginas 23 e 31.

DENISSEN, J. J. A. et al. Development and validation of a dutch translation of the big five inventory (bfi). *Journal of Personality Assessment*, Routledge, v. 90, n. 2, p. 152–157, 2008. Citado na página 23.

DIEZ, D. M.; BARR, C.; CETINKAYA-RUNDEL, M. *OpenIntro Statistics*. 3º edição. ed. [S.l.]: Organização OpenIntro, 2015. Citado na página 57.

DOSUALDO, D.; REZENDE, S. *Análise da Precisão de Métodos de Regressão*. [S.l.], 2003. Citado na página 57.

EDMONDS, G. W. et al. Is character fate, or is there hope to change my personality yet? *Social and Personality Psychology Compass*, v. 2, n. 1, p. 399–413, 2008. Citado na página 20.

EYSENCK, H. J. The scientific study of personality: A note on the review. *British Journal of Statistical Psychology*, Wiley Online Library, v. 6, n. 1, p. 44–52, 1953. Citado na página 21.

EYSENCK, H. J. Personality, genetics, and behavior: Selected papers. *Journal of the History of the Behavioral Sciences*, v. 19, n. 3, p. 284–286, 1983. Citado na página 21.

FABIO, A. D. Decisional procrastination correlates: personality traits, self-esteem or perception of cognitive failure? *International Journal for Educational and Vocational Guidance*, v. 6, n. 2, p. 109–122, Jun 2006. Citado na página 35.

FALCONER, A.; ADRAGNA, S. Personality types and persistence in doctoral students: A mixed-methods study. *Psychology and Behavioral Science International Journal*, v. 2, 03 2017. Citado na página 35.



FEIST, J.; FEIST, G. J.; TOMI-ANN, R. *Teorias da Personalidade*. 8. ed. [S.l.]: AMGH Editora, 2015. ISBN 0073532193. Citado na página 17.

FEYTER, T. D. et al. Unraveling the impact of the big five personality traits on academic performance: The moderating and mediating effects of self-efficacy and academic motivation. *Learning and Individual Differences*, v. 22, n. 4, p. 439 – 448, 2012. ISSN 1041-6080. Citado na página 13.

FOSSATI, A. et al. The big five inventory (bfi). *European Journal of Psychological Assessment*, v. 27, n. 1, p. 50–58, 2011. Citado na página 23.

FRUYT, F. de; MERVIELDE, I. Personality and interests as predictors of educational streaming and achievement. *European Journal of Personality*, v. 10, n. 5, p. 405–425, 1996. Citado na página 35.

FURNHAM, A.; MONSEN, J. Personality traits and intelligence predict academic school grades. *Learning and Individual Differences*, v. 19, n. 1, p. 28 – 33, 2009. ISSN 1041-6080. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1041608008000204>>. Citado na página 35.

GHORBANI, F.; MONTAZER, G. A. E-learners personality identifying using their network behaviors. *Computers in Human Behavior*, v. 51, p. 42 – 52, 2015. ISSN 0747-5632. Citado 4 vezes nas páginas 9, 27, 31 e 32.

GOLLAPUDI, S. *Practical Machine Learning*. 1. ed. [S.l.]: Packt Publishing, 2016. ISBN 9781784399689. Citado na página 29.

GROVES, R. M. *Survey Errors and Survey Costs*. 1. ed. [S.l.]: John Wiley & Sons, 1989. ISBN 0-471-61171-9. Citado na página 13.

HALAWA, M. S.; SHEHAB, M. E.; HAMED, E. M. R. Predicting student personality based on a data-driven model from student behavior on lms and social networks. In: *2015 Fifth International Conference on Digital Information Processing and Communications (ICDIPC)*. [S.l.: s.n.], 2015. p. 294–299. Citado 3 vezes nas páginas 9, 27 e 28.

HALL, C. S.; LINDZEY, G.; CAMPBELL, J. B. *Teorias da Personalidade*. 4. ed. [S.l.]: Artmed, 2007. ISBN 9788536307893. Citado 2 vezes nas páginas 18 e 19.

HALL, M. et al. The weka data mining software: An update. *SIGKDD Explor. Newsl.*, Association for Computing Machinery, New York, NY, USA, v. 11, n. 1, p. 10–18, nov. 2009. ISSN 1931-0145. Disponível em: <<https://doi.org/10.1145/1656274.1656278>>. Citado na página 48.

HOLMES, G.; HALL, M.; FRANK, E. Generating rule sets from model trees. In: *Advanced Topics in Artificial Intelligence*. [S.l.]: Springer Berlin Heidelberg, 1999. p. 1–12. Citado 2 vezes nas páginas 50 e 51.

JACKSON, J. J. et al. Your friends know how long you will live: A 75-year study of peer-rated personality traits. *Psychological Science*, v. 26, n. 3, p. 335–340, 2015. Citado na página 17.

JAMES, G. et al. *An Introduction to Statistical Learning*. 1. ed. [S.l.]: Springer, 2013. ISBN 9781461471370. Citado na página 46.

JOHN, O. P.; DONAHUE, E. M.; KENTLE, R. L. The big-five inventory. *Journal of Personality and Social Psychology*, 01 1991. Citado na página 23.

JOHN, O. P.; DONAHUE, E. M.; KENTLE, R. L. *The Big Five Inventory—Versions 4a and 54*. Berkeley, CA, 1991. Disponível em: <<https://www.ocf.berkeley.edu/~johnlab/index.htm>>. Citado na página 23.

JUDGE, T. A. et al. The big five personality traits, general mental ability, and career success across the life span. *Personnel Psychology*, v. 52, n. 3, p. 621–652, 1999. Citado 2 vezes nas páginas 13 e 17.

KOLLER, D. et al. Retention and intention in massive open online courses: In depth. *Educause Review*, v. 48, p. 62–63, 01 2013. Citado na página 30.

KOMARRAJU, M.; KARAU, S. J.; SCHMECK, R. R. Role of the big five personality traits in predicting college students' academic motivation and achievement. *Learning and Individual Differences*, v. 19, n. 1, p. 47–52, 2009. ISSN 1041-6080. Citado na página 13.

KOTSIANTIS, S.; PIERRAKEAS, C.; PINTELAS, P. Predicting student's performance in distance learning using machine learning techniques. *Applied Artificial Intelligence*, v. 18, n. 5, p. 411–426, 2004. Citado na página 25.

LAURIOLA, M.; LEVIN, I. Personality traits and risky decision-making in a controlled experimental task: An exploratory study. *Personality and Individual Differences*, v. 31, p. 215–226, 07 2001. Citado na página 17.

LEE, M.-C. Explaining and predicting users' continuance intention toward e-learning: An extension of the expectation–confirmation model. *Computers & Education*, Elsevier, v. 54, n. 2, p. 506–516, 2010. Citado na página 25.

LOUNSBURY, J. W. et al. Intelligence, “big five” personality traits, and work drive as predictors of course grade. *Personality and Individual Differences*, v. 35, n. 6, p. 1231 – 1239, 2003. ISSN 0191-8869. Citado na página 24.

Machine Learning Group at the University of Waikato. *Weka 3: Data Mining Software in Java*. 2019. <<http://www.cs.waikato.ac.nz/ml>>. Acessado em 10/12/2019. Citado na página 48.

MANFREDA, K. L. et al. Web surveys versus other survey modes: A meta-analysis comparing response rates. *International Journal of Market Research*, v. 50, p. 79–104, 01 2008. Citado na página 13.

MCCRAE, R. R. et al. Cross-cultural assessment of the five-factor model: The revised neo personality inventory. *Journal of Cross-Cultural Psychology*, v. 29, n. 1, p. 171–188, 1998. Citado na página 23.

MCCRAE, R. R.; COSTA, P. T. Set like plaster? evidence for the stability of adult personality. p. 21–40, 01 1994. Citado na página 20.

MCCRAE, R. R.; COSTA, P. T. The stability of personality: Observations and evaluations. *Current Directions in Psychological Science*, [Association for Psychological Science, Sage Publications, Inc.], v. 3, n. 6, p. 173–175, 1994. Citado na página 20.

MONTGOMERY, D. *Design and Analysis of Experiments, 8th Edition*. [S.l.]: John Wiley & Sons, Incorporated, 2012. ISBN 9781118214718. Citado na página 57.

NETTLE, D. The evolution of personality variation in humans and other animals. *American Psychologist*, v. 61, n. 6, p. 622–631, 09 2006. Citado na página 20.

OSHIO, A. et al. Resilience and big five personality traits: A meta-analysis. *Personality and Individual Differences*, v. 127, p. 54–60, 2018. ISSN 0191-8869. Citado 2 vezes nas páginas 13 e 17.

OZER, D. J.; BENET-MARTÍNEZ, V. Personality and the prediction of consequential outcomes. *Annual Review of Psychology*, v. 57, n. 1, p. 401–421, 2006. Citado 2 vezes nas páginas 13 e 17.

PALMA, M. T. M. M. *A Prossecução dos Estudos: O Papel da Personalidade na Tomada de Decisão*. Lisboa, Portugal: [s.n.], 2006. Citado na página 23.

PASQUALI, L. *Os Tipos Humanos: A Teoria Da Personalidade*. 1. ed. [S.l.]: Vozes, 2003. ISBN 8532628621. Citado na página 18.

PAUNONEN, S. V.; ASHTON, M. C. On the prediction of academic performance with personality traits: A replication study. *Journal of Research in Personality*, v. 47, n. 6, p. 778–781, 2013. ISSN 0092-6566. Citado na página 13.

POROPAT, A. A meta-analysis of the five-factor model of personality and academic performance. *Psychological bulletin*, v. 135, p. 322–38, 04 2009. Citado 2 vezes nas páginas 24 e 35.

RAAD, B. D.; SCHOUWENBURG, H. C. Personality in learning and education: a review. *European Journal of Personality*, Wiley Online Library, v. 10, n. 5, p. 303–336, 12 1996. ISSN 1099-0984. Disponível em: <[https://doi.org/10.1002/\(SICI\)1099-0984\(199612\)10:5<303::AID-PER262>3.0.CO;2-2](https://doi.org/10.1002/(SICI)1099-0984(199612)10:5<303::AID-PER262>3.0.CO;2-2)>. Citado 2 vezes nas páginas 24 e 35.

RICHARDSON, M.; ABRAHAM, C.; BOND, R. Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin*, v. 138, p. 353–87, 03 2012. Citado na página 24.

ROLFHUS, E. L.; ACKERMAN, P. L. Assessing individual differences in knowledge: Knowledge, intelligence, and related traits. *Journal of Educational Psychology*, v. 91, p. 511–526, 1999. Citado 2 vezes nas páginas 24 e 35.

ROLLAND, J.-P. The five-factor model of personality across cultures. In: \_\_\_\_\_. [S.l.: s.n.], 2002. (International and Cultural Psychology), p. 7–28. ISBN HB: 0-306-47354-2 PN: 0-306-47355-0. Citado na página 23.

ROLLAND2002. *The Five-Factor Model of Personality Across Cultures*. [S.l.]: Springer US, 2002. (International and Cultural Psychology). ISBN 9780306473548. Citado na página 38.

ROMERO, C. et al. Predicting students' final performance from participation in on-line discussion forums. *Computers & Education*, v. 68, p. 458 – 472, 2013. Citado na página 25.

ROMERO, C.; VENTURA, S. Educational data mining: A survey from 1995 to 2005. *Expert Systems with Applications*, v. 33, n. 1, p. 135–146, 2007. Citado na página 25.

ROMERO, C.; VENTURA, S. Educational data mining: A review of the state of the art. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, v. 40, p. 601 – 618, 12 2010. Citado 2 vezes nas páginas 24 e 25.

ROTHMANN, S.; COETZER, E. The big five personality dimensions and job performance. *South African Journal of Industrial Psychology*, v. 29, p. 68–74, 10 2003. Citado na página 17.

SATCHELL, L. et al. Ruminating on the nature of intelligence: Personality predicts implicit theories and educational persistence. *Personality and Individual Differences*, v. 113, p. 109 – 114, 2017. ISSN 0191-8869. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0191886917301952>>. Citado na página 35.

SCHMITT, D. P. et al. The geographic distribution of big five personality traits: Patterns and profiles of human self-description across 56 nations. *Journal of Cross-Cultural Psychology*, v. 38, n. 2, p. 173–212, 2007. Disponível em: <<https://doi.org/10.1177/0022022106297299>>. Citado 3 vezes nas páginas 23, 37 e 38.

Silva, R. M.; Yamakami, A.; Almeida, T. A. An analysis of machine learning methods for spam host detection. In: *2012 11th International Conference on Machine Learning and Applications*. [S.l.: s.n.], 2012. v. 2, p. 227–232. ISSN null. Citado na página 50.

SRIVASTAVA, S. M.; JOHN, O. P. The big-five trait taxonomy: History, measurement, and theoretical perspectives. In: PERVIN; JOHN, O. P. (Ed.). *Handbook of personality: Theory and research*. [S.l.]: Guilford Press., 1999. cap. 4, p. 102–138. Citado 3 vezes nas páginas 21, 23 e 38.

STEEL, P.; BROTHEN, T.; WAMBACH, C. Procrastination and personality, performance, and mood. *Personality and Individual Differences*, v. 30, n. 1, p. 95 – 106, 2001. ISSN 0191-8869. Citado na página 24.

VEDEL, A. The big five and tertiary academic performance: A systematic review and meta-analysis. *Personality and Individual Differences*, v. 71, p. 66 – 76, 2014. ISSN 0191-8869. Citado 2 vezes nas páginas 24 e 35.

VERMETTEN, Y. J.; LODEWIJKS, H. G.; VERMUNT, J. D. The role of personality traits and goal orientations in strategy use. *Contemporary Educational Psychology*, v. 26, n. 2, p. 149 – 170, 2001. ISSN 0361-476X. Citado na página 24.

WITTEN, I. H.; FRANK, E.; HALL, M. A. *Data Mining: Practical Machine Learning Tools and Techniques*. 3. ed. [S.l.]: Elsevier, 2011. ISBN 978-0-12-374856-0. Citado 3 vezes nas páginas 29, 47 e 49.

ZAMORANO, E. R. et al. Psychometric properties of the big five inventory in a mexican sample (propiedades psicométricas del inventario de los cinco grandes en una muestra mexicana). *Salud Mental*, v. 37, p. 491, 12 2014. Citado na página 23.

ZHENG, L. et al. Reliability and concurrent validation of the ipip big-five factor markers in china: Consistencies in factor structure between internet-obtained heterosexual and homosexual samples. *Personality and Individual Differences*, v. 45, n. 7, p. 649 – 654, 2008. ISSN 0191-8869. Citado na página 23.

## ANEXO A – QUESTIONÁRIO INVENTÁRIO DOS CINCO GRANDES FATORES DE PERSONALIDADE

**INSTRUÇÕES.** A seguir encontram-se algumas características que podem ou não lhe dizer respeito. Por favor, escolha um dos números na escala abaixo que melhor expresse sua opinião em relação a você mesmo e anote no espaço ao lado de cada afirmação. Vale ressaltar que não existem respostas certas ou erradas. Utilize a seguinte escala de resposta:

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Discordo totalmente	Discordo em parte	Nem concordo nem discordo	Concordo em parte	Concordo totalmente

**Eu me vejo como alguém que...**

- \_\_\_01. É conversador, comunicativo.
- \_\_\_02. Às vezes é frio e distante.
- \_\_\_03. Tende a ser crítico com os outros.
- \_\_\_04. É minucioso, detalhista no trabalho.
- \_\_\_05. É assertivo, não teme expressar o que sente.
- \_\_\_06. Insiste até concluir a tarefa ou o trabalho.
- \_\_\_07. É depressivo, triste.
- \_\_\_08. Gosta de cooperar com os outros.
- \_\_\_09. É original, tem sempre novas idéias.
- \_\_\_10. É temperamental, muda de humor facilmente.
- \_\_\_11. É inventivo, criativo.
- \_\_\_12. É reservado.
- \_\_\_13. Valoriza o artístico, o estético.
- \_\_\_14. É emocionalmente estável, não se altera facilmente.
- \_\_\_15. É prestativo e ajuda os outros.
- \_\_\_16. É, às vezes, tímido, inibido.
- \_\_\_17. Pode ser um tanto descuidado.
- \_\_\_18. É amável, tem consideração pelos outros.
- \_\_\_19. Tende a ser preguiçoso.
- \_\_\_20. Faz as coisas com eficiência.
- \_\_\_21. É relaxado, controla bem o estresse.
- \_\_\_22. É facilmente distraído.
- \_\_\_23. Mantém-se calmo nas situações tensas.
- \_\_\_24. Prefere trabalho rotineiro.
- \_\_\_25. É curioso sobre muitas coisas diferentes.
- \_\_\_26. É sociável, extrovertido.
- \_\_\_27. É geralmente confiável.
- \_\_\_28. É, às vezes, rude (grosseiro) com os outros.
- \_\_\_29. É cheio de energia.
- \_\_\_30. Começa discussões, disputas com os outros.
- \_\_\_31. É um trabalhador de confiança.
- \_\_\_32. Faz planos e os segue a risca.
- \_\_\_33. Tem uma imaginação fértil.
- \_\_\_34. Fica tenso com frequência.
- \_\_\_35. É engenhoso, alguém que gosta de analisar profundamente as coisas.
- \_\_\_36. Fica nervoso facilmente.
- \_\_\_37. Gera muito entusiasmo.
- \_\_\_38. Tende a ser desorganizado.
- \_\_\_39. Gosta de refletir, brincar com as idéias.
- \_\_\_40. Tem capacidade de perdoar, perdoa fácil.
- \_\_\_41. Preocupa-se muito com tudo.
- \_\_\_42. Tende a ser quieto, calado.
- \_\_\_43. Tem poucos interesses artísticos.
- \_\_\_44. É sofisticado em artes, música ou literatura.



## ANEXO B – AUTORIZAÇÃO INSTITUCIONAL PARA REALIZAÇÃO DA PESQUISA – CIED



### *Autorização Institucional*

Eu, **Diego de Oliveira Souza**, responsável pela Coordenadoria Institucional de Educação a Distância (CIED) da Universidade Federal de Alagoas, declaro que fui informado dos objetivos da pesquisa “Detecção Automática de Tipos de Personalidades de Alunos em Ambientes de Ensino a Distância”, a ser executada pelo Sr. David Medeiros Batinga, orientado pelos professores Dr. Ranilson Oscar Araújo Paiva e Dr. Jorge Artur Peçanha de Miranda Coelho e concordo em autorizar a execução da mesma nesta instituição. Caso necessário, a qualquer momento como instituição CO-PARTICIPANTE desta pesquisa poderemos revogar esta autorização, se comprovada atividades que causem algum prejuízo à esta instituição ou ainda, a qualquer dado que comprometa o sigilo da participação dos integrantes desta instituição. Declaro também, que não recebemos qualquer pagamento por esta autorização bem como os participantes não receberão qualquer tipo de pagamento.

Estou ciente de que o projeto terá seu início somente após a aprovação pelo Comitê de Ética em Pesquisa, mediante parecer ético consubstanciado.

Declaro que conheço a Resolução 510/2016 do Conselho Nacional de Saúde e que seus preceitos serão seguidos.

Maceió, 11 de Setembro de 2018.

 David Medeiros Batinga Pesquisador	 Diego de Oliveira Souza <small>Universidade Federal de Alagoas</small> Diego de Oliveira Souza Responsável pela CIED
--	---





## ANEXO C – PARECER CONSUBSTANCIADO DO CEP

UNIVERSIDADE FEDERAL DE  
ALAGOAS



### PARECER CONSUBSTANCIADO DO CEP

#### DADOS DO PROJETO DE PESQUISA

**Título da Pesquisa:** Detecção Automática de Tipos de Personalidades de Alunos em Ambientes de Ensino a Distância

**Pesquisador:** DAVID MEDEIROS BATINGA

**Área Temática:**

**Versão:** 2

**CAAE:** 99895118.2.0000.5013

**Instituição Proponente:** Universidade Federal de Alagoas

**Patrocinador Principal:** Financiamento Próprio

#### DADOS DO PARECER

**Número do Parecer:** 3.082.482

#### Apresentação do Projeto:

Pesquisas indicam que os traços de personalidade influenciam vários aspectos da vida cotidiana e exercem um importante papel para descrever padrões de comportamento que uma pessoa demonstra, quando confrontada com diversos estímulos e situações; incluindo, entre outros, sua percepção acerca da dedicação e realização nos estudos e no trabalho (JUDGE et al., 1999); sua facilidade em estabelecer relacionamentos interpessoais, resiliência a dificuldades e desafios (OSHIO et al., 2018), e concepção sobre a sua autoestima e felicidade (OZER; BENET-MARTÍNEZ, 2006).

O estudo da personalidade passou por um processo de estruturação e formalização na área da Psicologia a partir da década de 1930. Um dos marcos desse processo foi a publicação da obra de ALLPORT em 1938, que definiu traço de personalidade como um sistema neuropsíquico, intrínseco a cada indivíduo, capaz de receber e converter muitos estímulos funcionalmente equivalentes em padrões de comportamento congruentes. Em outras palavras, os traços de personalidade são predisposições a responder igualmente ou de modo semelhante a determinados estímulos.

A presente pesquisa tem como foco investigar o conceito de traços de personalidade no âmbito da educação, mais precisamente ao âmbito da modalidade de ensino à distância (EAD). Trabalhos anteriores demonstraram que o conceito de personalidade é de relevância prática para a área educacional, tendo sido provadas correlações entre os traços de

**Endereço:** Av. Lourival Melo Mota, s/n - Campus A . C. Simões,

**Bairro:** Cidade Universitária

**CEP:** 57.072-900

**UF:** AL

**Município:** MACEIO

**Telefone:** (82)3214-1041

**E-mail:** comitedeeticaufal@gmail.com

Continuação do Parecer: 3.082.482

personalidade com temas como: modelo e agrupamento de alunos, desempenho acadêmico (FEYTER et al., 2012; PAUNONEN; ASHTON, 2013), nível de engajamento e dedicação dos alunos (KOMARRAJU; KARAU; SCHMECK, 2009).

A análise de traços de personalidade de indivíduos é realizada, tradicionalmente, por meio da aplicação de questionários psicométricos (ANDRADE, 2008). Em se tratando da modalidade EAD, a utilização dessa abordagem se torna complexa e dispendiosa, tendo em vista que os cursos à distância geralmente apresentam um número maior de alunos, quando comparados aos cursos presenciais. Para contornar essa dificuldade, há a possibilidade de se adotar o envio de questionários por meio eletrônico, mas, em geral, o percentual de retorno para esse tipo de contato é baixo. Um outro problema associado ao uso de questionários (físicos) é que há margem para o preenchimento errado (deliberado ou não) dos questionários; e ainda permite o envio de respostas influenciadas pelo interesse em ampliar ou em omitir características do respondente. Esta pesquisa pretende definir um modelo computacional, apoiado pela teoria de traços de personalidade, que apoie a classificação automática do tipo de personalidade dos alunos por meio de técnicas de mineração de dados educacionais. Segundo a proposta, essa análise será baseada exclusivamente nos registros dos hábitos, ações e interações realizados pelos alunos durante as atividades pedagógicas realizadas num Ambiente Virtual de Aprendizagem (AVA). E que quando implementado, este modelo sirva como ferramenta prática para que professores, tutores e equipe pedagógica possam identificar os traços de personalidade dos alunos, dispensando totalmente a necessidade de aplicação de questionários psicométricos.

**Objetivo da Pesquisa:**

Objetivo primário

Propor um modelo computacional que permita a identificação automática da personalidade dos estudantes.

Objetivo secundário

Identificar os traços de personalidade predominantes dos alunos de ensino à distância da UFAL.

**Avaliação dos Riscos e Benefícios:**

Riscos

Os incômodos e possíveis riscos à saúde física e/ou mental dos participantes são que o preenchimento do questionário desta pesquisa poderá expô-los a riscos mínimos

**Endereço:** Av. Lourival Melo Mota, s/n - Campus A . C. Simões,**Bairro:** Cidade Universitária**CEP:** 57.072-900**UF:** AL**Município:** MACEIO**Telefone:** (82)3214-1041**E-mail:** comitedeeticaufal@gmail.com

## UNIVERSIDADE FEDERAL DE ALAGOAS



Continuação do Parecer: 3.082.482

como cansaço, desconforto pelo tempo gasto no preenchimento do questionário e possível ao receber o resultado do questionário aplicado e ao tomar consciência de que a sua atuação está em processo de análise estatística por esta pesquisa. Se isto ocorrer, o participante poderá interromper o preenchimento dos instrumentos e retorná-los posteriormente, se assim o desejar.

Aos participantes que sentirem algum desconforto psicológico, em decorrência da participação na pesquisa, será providenciada assistência psicológica adequada. O atendimento psicológico será realizado, mediante agendamento, pela equipe de psicólogos do Serviço de Psicologia Aplicada (SPA) do Instituto de Psicologia da Universidade Federal de Alagoas (UFAL), conforme o termo anexado a esta proposta de pesquisa (Anexo C).

### Benefícios

O modelo computacional proposto servirá como ferramenta prática para que professores e equipe pedagógica conheçam melhor os traços de personalidade dos alunos; viabilizando, assim, a implementação de melhorias nas intervenções pedagógicas personalizadas, que objetivem o incremento na aprendizagem.

### Comentários e Considerações sobre a Pesquisa:

A pesquisa segue uma metodologia do tipo descritiva-exploratória com o objetivo principal de desenvolver um modelo computacional que permita a classificação automática do tipo de personalidade de alunos de cursos de educação à distância.

O modelo será desenvolvido segundo um processo organizado em quatro etapas:

1. Num primeiro momento, um banco de dados será construído, no qual serão armazenados os dados dos alunos oriundos das respostas de um questionário psicométrico e dos registros de log das interações efetuadas pelos alunos no ambiente virtual de aprendizagem. Em relação ao questionário, essa pesquisa adotará o modelo de questionário Big Five Inventory (Inventário dos Cinco Grandes Fatores de Personalidade – IGFP-5) adaptado e validado para o Brasil. O IGFP-5 é uma medida de auto-relato breve, composta por 44 itens e designada a avaliar dimensões da personalidade baseada no modelo dos Cinco Grandes Fatores da Personalidade (ANDRADE, 2008), conforme modelo anexado a esta proposta de projeto de pesquisa (Anexo A). O instrumento IGFP-5 foi escolhido para esta pesquisa porque demonstrou resultados consistentes em vários estudos anteriores e porque apresenta nível de confiabilidade equivalente a outros instrumentos da mesma classe, mesmo expondo

**Endereço:** Av. Lourival Melo Mota, s/n - Campus A . C. Simões,

**Bairro:** Cidade Universitária

**CEP:** 57.072-900

**UF:** AL

**Município:** MACEIO

**Telefone:** (82)3214-1041

**E-mail:** comitedeeticaufal@gmail.com

Continuação do Parecer: 3.082.482

um número menor de itens (questões) aos respondentes (SRIVASTAVA; JOHN, 1999).

A estrutura penta-dimensional do instrumento foi validada e se mostrou robusta na maioria das culturas e regiões do mundo (MCCRAE et al., 1998; MCCRAE; ALLIK, 2002; SCHMITT et al., 2007) (MCCRAE; ALLIK, 2002 apud ANDRADE, 2008), tendo isso também comprovada sua adequação à realidade brasileira (ANDRADE, 2008).

Sustentado em décadas de pesquisa analítica fatorial, o instrumento pressupõe avaliar as cinco grandes dimensões da personalidade humana:

**Abertura (Openness):** Indivíduos com alta pontuação nesta dimensão, geralmente, são francos, imaginativos, espirituosos, originais e artísticos. Por outro lado, indivíduos com baixa pontuação nesta dimensão são superficiais, comuns ou simples.

**Conscienciosidade (Conscientiousness):** Nesse fator é característico o controle de impulsos, bem como comportamentos direcionados a um objetivo específico, que podem facilitar a execução de obrigações e deveres. Indivíduos conscienciosos são geralmente cautelosos, dignos de confiança, organizados e responsáveis. Por outro lado, indivíduos com baixos escores nessa dimensão tendem a ser descuidados, desordenados e pouco-confiáveis.

**Extroversão (Extraversion):** Também é chamada de “Expansão”; indivíduos extrovertidos tendem a serem ativos, entusiasmados, dominantes, sociáveis e eloquentes ou falantes. Por outro lado, indivíduos introvertidos tendem a ser retraídos, submissos e quietos.

**Amabilidade (Agreeableness):** Também é chamado comumente de “Agradabilidade” ou “Sociabilidade”. Indivíduos com altas pontuações nesse traço são agradáveis, amáveis, cooperativos e afetuosos. Indivíduos com uma classificação baixa nessa dimensão podem ser frios e indelicados. A referida dimensão caracteriza-se por uma orientação em direção aos demais, incluindo traços como altruísmo, confiança e modéstia.

**Neuroticismo (Neuroticism):** Também é chamado de “Instabilidade emocional”. Indivíduos neuróticos são geralmente nervosos, altamente sensíveis, tensos e preocupados. Por outro lado, indivíduos emocionalmente estáveis são calmos e satisfeitos. O traço Neuroticismo contrasta estabilidade emocional com afetos negativos, incluindo ansiedade, tristeza, irritabilidade e tensão nervosa.

**Endereço:** Av. Lourival Melo Mota, s/n - Campus A . C. Simões,

**Bairro:** Cidade Universitária

**CEP:** 57.072-900

**UF:** AL

**Município:** MACEIO

**Telefone:** (82)3214-1041

**E-mail:** comitedeeticaufal@gmail.com

## UNIVERSIDADE FEDERAL DE ALAGOAS



Continuação do Parecer: 3.082.482

O questionário será aplicado à uma amostra de 286 alunos dos cursos de educação à distância da UFAL, considerando uma população de 1112 alunos de cursos da modalidade de ensino à distância, com matrículas ativas e que cursaram (ou estão cursando) pelo menos uma disciplina no decorrer do ano 2018. Para o cálculo da amostragem foi considerado um nível de confiança de 95% e uma margem de erro de 5%.2. Após a aplicação dos questionários, os dados preenchidos serão transcritos para uma planilha eletrônica. Os originais dos questionários respondidos serão guardados em local seguro e protegido, sob os cuidados do responsável da pesquisa. Na sequência, será obtida, junto à CIED, cópia dos logs das interações realizadas pelos participantes do experimento durante as atividades pedagógicas efetuadas no ambiente de ensino à distância da UFAL (plataforma Moodle). Neste instante, os dados serão anonimizados. Os nomes dos participantes serão substituídos por um número atribuído aleatoriamente tanto na planilha das respostas do questionário, quanto nos logs do Moodle. A partir deste ponto, os nomes dos participantes não mais constarão nos dados manipulados pelas próximas fases do experimento.

3. No estágio seguinte, será realizado um experimento para seleção do algoritmo de mineração de dados para classificação a ser utilizado na construção do modelo.

4. Finalmente, na etapa de desenvolvimento do modelo, será construído e validado o modelo para classificação dos traços de personalidade dos estudantes. Os resultados obtidos a partir das respostas do questionário serão essenciais para o cumprimento dessa etapa, viabilizando a definição e validação do modelo computacional proposto.

### **Considerações sobre os Termos de apresentação obrigatória:**

Apresentados satisfatoriamente.

A primeira lista de pendências (Esclarecer os procedimentos metodológicos e desfechos que se pretende com o formulário de pesquisa,

bem como fundamentar a escolha do instrumento de pesquisa;

- Categorizar os tipos de personalidade de acordo com o instrumento de pesquisa;
- Descrever os procedimentos no TCLE, pois o participante de pesquisa precisa compreender com detalhes como participará do projeto;
- Explicar a análise de dados;
- Descrever como vai prestar assistência no caso de um diagnóstico de um perfil psicológico.

**Endereço:** Av. Lourival Melo Mota, s/n - Campus A . C. Simões,

**Bairro:** Cidade Universitária

**CEP:** 57.072-900

**UF:** AL

**Município:** MACEIO

**Telefone:** (82)3214-1041

**E-mail:** comitedeeticaufal@gmail.com

UNIVERSIDADE FEDERAL DE  
ALAGOAS



Continuação do Parecer: 3.082.482

- Descrever como fará a anonimização dos dados; O instrumento foi validado transculturalmente?
- Esclarecer amostragem e cálculo amostral.) foi respondida satisfatoriamente.

**Recomendações:**

Presença de psicólogo na equipe de pesquisa.

**Conclusões ou Pendências e Lista de Inadequações:**

Protocolo atende a Resolução 510/16.

**Considerações Finais a critério do CEP:**

Protocolo Aprovado

Prezado (a) Pesquisador (a), lembre-se que, segundo a Res. CNS 466/12 e sua complementar 510/2016:

O participante da pesquisa tem a liberdade de recusar-se a participar ou de retirar seu consentimento em qualquer fase da pesquisa, sem penalização alguma e sem prejuízo ao seu cuidado e deve receber cópia do TCLE, na íntegra, por ele assinado, a não ser em estudo com autorização de declínio;

V.S<sup>a</sup>. deve desenvolver a pesquisa conforme delineada no protocolo aprovado e descontinuar o estudo somente após análise das razões da descontinuidade por este CEP, exceto quando perceber risco ou dano não previsto ao sujeito participante ou quando constatar a superioridade de regime oferecido a um dos grupos da pesquisa que requeiram ação imediata;

O CEP deve ser imediatamente informado de todos os fatos relevantes que alterem o curso normal do estudo. É responsabilidade do pesquisador assegurar medidas imediatas adequadas a evento adverso ocorrido e enviar notificação a este CEP e, em casos pertinentes, à ANVISA;

Eventuais modificações ou emendas ao protocolo devem ser apresentadas ao CEP de forma clara e sucinta, identificando a parte do protocolo a ser modificada e suas justificativas. Em caso de projetos do Grupo I ou II apresentados anteriormente à ANVISA, o pesquisador ou patrocinador deve enviá-las também à mesma, junto com o parecer aprovatório do CEP, para serem juntadas ao protocolo inicial;

Seus relatórios parciais e final devem ser apresentados a este CEP, inicialmente após o prazo determinado no seu cronograma e ao término do estudo. A falta de envio de, pelo menos, o relatório final da pesquisa implicará em não recebimento de um próximo protocolo de pesquisa de vossa autoria.

O cronograma previsto para a pesquisa será executado caso o projeto seja APROVADO pelo Sistema CEP/CONEP, conforme Carta Circular nº. 061/2012/CONEP/CNS/GB/MS (Brasília-DF, 04

**Endereço:** Av. Lourival Melo Mota, s/n - Campus A . C. Simões,

**Bairro:** Cidade Universitária

**CEP:** 57.072-900

**UF:** AL

**Município:** MACEIO

**Telefone:** (82)3214-1041

**E-mail:** comitedeeticaufal@gmail.com

UNIVERSIDADE FEDERAL DE  
ALAGOAS



Continuação do Parecer: 3.082.482

de maio de 2012).

**Este parecer foi elaborado baseado nos documentos abaixo relacionados:**

Tipo Documento	Arquivo	Postagem	Autor	Situação
Informações Básicas do Projeto	PB_INFORMAÇÕES_BÁSICAS_DO_PROJETO_1108223.pdf	02/11/2018 20:44:04		Aceite
Outros	Atestado_assistencia_psicologica.pdf	02/11/2018 20:43:00	DAVID MEDEIROS BATINGA	Aceite
Projeto Detalhado / Brochura Investigador	Projeto_de_pesquisa_v2.pdf	02/11/2018 20:33:42	DAVID MEDEIROS BATINGA	Aceite
TCLE / Termos de Assentimento / Justificativa de Ausência	TCLEv2.pdf	02/11/2018 20:33:12	DAVID MEDEIROS BATINGA	Aceite
Outros	Questionario_participantes.pdf	01/10/2018 14:12:00	DAVID MEDEIROS BATINGA	Aceite
Declaração de Pesquisadores	Declaracao_de_cumprimento_das_resolucoes_422_2012_e_510_2016.pdf	01/10/2018 14:07:27	DAVID MEDEIROS BATINGA	Aceite
Folha de Rosto	Folha_de_rosto_projeto_pesquisa_David_Medeiros_Batinga.pdf	28/09/2018 11:30:27	DAVID MEDEIROS BATINGA	Aceite
Declaração de Instituição e Infraestrutura	Autorizacao_Institucional_CIED_Coleta_de_dados.pdf	12/09/2018 18:34:18	DAVID MEDEIROS BATINGA	Aceite

**Situação do Parecer:**

Aprovado

**Necessita Apreciação da CONEP:**

Não

MACEIO, 13 de Dezembro de 2018

Assinado por:  
**Luciana Santana**  
(Coordenador(a))

**Endereço:** Av. Lourival Melo Mota, s/n - Campus A . C. Simões,

**Bairro:** Cidade Universitária

**CEP:** 57.072-900

**UF:** AL

**Município:** MACEIO

**Telefone:** (82)3214-1041

**E-mail:** comitedeeticaufal@gmail.com