FEDERAL UNIVERSITY OF ALAGOAS - UFAL
COMPUTING INSTITUTE - IC
GRADUATE PROGRAM IN COMPUTATIONAL MODELLING OF KNOWLEDGE - PPGMCC

JOSMÁRIO DE ALBUQUERQUE SILVA

A supervised learning approach to detect gender stereotype in online educational technologies

Maceió – Alagoas – Brazil

June of 2019

JOSMÁRIO DE ALBUQUERQUE SILVA

A supervised learning approach to detect gender stereotype in online educational technologies

A thesis presented to the Graduate Program of Computational Modeling of Knowledge from the Computing Institute at the Federal University of Alagoas in partial fulfillment of requirements for the Master of Science Degree.

Supervisor: Prof. Dr. Ig I. Bittencourt
Co-supervisor: Prof. Dr. Jorge. A. P. M. Coelho

Maceió – Alagoas – Brazil

June of 2019

Catalogação na fonte
Universidade Federal de Alagoas
Biblioteca Central
Divisão de Tratamento Técnico
Bibliotecário: Marcelino de Carvalho

S586s Silva, Josmário de Albuquerque.
A supervised learning approach to detect gender stereotype in online educational technologies / Josmário de Albuquerque Silva. – 2019.
80 f. : il.

Orientador: Ig Ibert Bittencourt Santana Pinto.
Coorientador: Jorge Artur Peçanha de Miranda Coelho.

Dissertação (mestrado em Modelagem Computacional de Conhecimento) – Universidade Federal de Alagoas. Instituto de Computação. Maceió, 2019.

Bibliografia: f. 59-72.
Apêndices: f. 73-80.

1. Tecnologia educacional. 2. Identidade de gênero na educação. 3. Estereótipos (Psicologia). I. Título.

CDU: 004:37

Folha de Aprovação

Josmário de Albuquerque Silva

A supervised learning approach to detect gender stereotype in online educational technologies

Dissertação submetida ao corpo docente do Programa de Pós-Graduação em Modelagem Computacional de Conhecimento da Universidade Federal de Alagoas e aprovada em 21 de junho de 2019.

Prof. Dr. Ig Ibert Bittencourt Santana Pinto
Instituto de Computação - UFAL
Orientador

Prof. Dr. Jorge Artur Peçanha de Miranda Coelho
Faculdade de Medicina - UFAL
Co-orientador

Banca Examinadora:

Leonardo Marques

Prof. Dr. Leonardo Brandão Marques
Centro de Educação- UFAL
Examinador Interno

Profa. Dra. Sheyla Christine Santos Fernandes
Instituto de Psicologia - UFAL
Examinador Externo

Wayne Holmes

Prof. Dr. Wayne Holmes
Open University - UK
Examinador Externo

To my family, friends,
and everyone who contributed
to this work.

ACKNOWLEDGEMENTS

By the end of this journey, I am very grateful for those who were always assisting me along the way. First, I do very appreciate the support of my advisors Prof. Dr. Ig Bittencourt and Prof. Dr. Jorge Coelho from the Federal University of Alagoas at the Computing Institute and College of Health, respectively; they were always available to assist me and guide me along this important phase of my life. Second, I thank the professors who evaluated this work and provided valuable feedback during the phases of qualification and final defense, for instance, Prof. Dr. Diego Matos, Prof. Dr. Sheyla Fernandes, Prof. Dr. Leonardo Marques, and Dr. Wayne Holmes. I would like to thank my family who has ever been very close to me even in the most difficult moments; without the support of my parents, João and Marlene, I would not have even start this program. I also appreciate the encouragement of my siblings, Aparecida and Lucas for being cheering for my success all the time. Beyond those, my aunts Ana Paula and Francisca have always been by my side along and I could not forget the importance of their help for me. In addition, I acknowledge my closest friends for always motivating me and helping me keep going even in the most challenging situations. I also thank the National Coordination for the Improvement of Higher Education Personnel (CAPES) who provided financial support through its National Program of Academic Cooperation (PROCAD). Also, the members from the Center for Excellence in Social Technologies (NEES) and the Applied Computing in Education Laboratory for their substantial contributions during preliminary presentations I performed. I am also grateful to all professors who contributed with important phases of my academic life by teaching me valuable professional and personal skills. Finally, I would like to thank my institution and everyone who directly and indirectly contributed to the development of this work.

"True equality means holding everyone accountable in the same way, regardless of race, gender, faith, ethnicity – or political ideology."

(Monica Crowley)

RESUMO

Tecnologia educacional (Edtech) tem impactado o modo como aprendemos e ensinamos, por exemplo, melhorando o engajamento de alunos, reforçando a colaboração, aumentando a retenção de aprendizado e ajudando professores a criar e distribuir novos conteúdos. No entanto, pesquisadores destacam que questões relacionadas à igualdade de gênero, como os estereótipos de gênero, precisam ser abordadas para promover contextos de aprendizagem plurais e inclusivos. De fato, descobertas recentes mostram que os estereótipos afetam vários aspectos no processo de aprendizagem, por exemplo, desempenho, engajamento, confiança, auto-imagem e ansiedade. No entanto, para abordar essas questões, precisamos antecipadamente descobrir se uma dada tecnologia educacional é estereotipada ou não. Diante desse cenário, propomos uma abordagem baseada em classificadores de aprendizagem supervisionada para detectar estereótipos de gênero em ambientes educacionais online. O método consiste na coleta de pistas situacionais de ameaça de estereótipos, isto é, conteúdos textuais e esquemas de cores de páginas Web para desenvolver e validar um modelo preditivo de aprendizagem de máquina. Além disso, a fim de validar o problema e reunir mais informações sobre o impacto dos estereótipos de gênero em tais contextos, realizamos uma revisão sistemática para destacar evidências e destacar as descobertas entre diferentes tipos de tecnologias educacionais e nos últimos 20 anos. A revisão também mostra abordagens metodológicas ao longe desses anos e as limitações de tais estudos. Em relação aos modelos preditivos, nossa abordagem mostrou uma alta precisão na previsão de ameaças de estereótipos de gênero em ambientes online. Também implementamos a abordagem e a aplicamos às páginas Web de universidades que se destacam no ranking mundial e os resultados sugerem a presença de estereótipos masculinos nas mesmas. Discutimos esses achados e apresentamos uma agenda de pesquisa para sublinhar os pontos que carecem de uma anteção especial na investigação de estereótipos de gênero e tecnologias educacionais.

Palavras-chaves: estudos de gênero; ameaça de estereótipo; educação online; aprendizagem supervisionada.

ABSTRACT

Educational Technology (Edtech) has impacted the way humans learn and teach, e.g., improving students' engagement, bolstering collaboration, increasing learning retention, and assisting teachers in creating and delivering new contents. However, researchers have highlighted that issues related to gender equality like gender stereotypes need to be addressed in order to promote plural and inclusive learning settings. In fact, recent findings show stereotypes have impacted several aspects in the learning process, e.g., performance, engagement, confidence, self-image, and anxiety. However, to address those issues, we require in advance to find out whether a given educational technology is stereotyped. Given that scenario, we propose an approach based on supervised learning classifiers to detect gender stereotypes in online educational environments. The method consisted of gathering situational cues of stereotype threat, i.e., textual contents and color schemes from web pages to develop and validate a machine learning predictive model. In addition, in order to validate the problem and gather more information about the impact of gender stereotypes in such settings, we primarily performed a systematic review to highlight evidence in the field and summarized the findings among different types of educational technologies and their implications in the last 20 years. The review also shows methodological approaches used along with those years and the limitations of such studies. Regarding predictive models, our approach showed a high precision on predicting gender stereotype threat in online settings. We also implemented the approach and applied it towards top-ranked universities' web pages and the results suggest the presence of male bias in such settings. We discuss those findings and present a research agenda to underline research points that should be concerned when investigating gender stereotypes and educational technologies.

Keywords: gender studies; stereotype threat; online education; supervised learning.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

ACM	Association for Computing Machinery
CART	Classification and Regression Tree
CSCL	Computer-Supported Collaborative Learning
EdTech	Educational Technology
FP	False Positive
HTML	HyperText Markup Language
IEEE	Institute of Electrical and Electronics Engineers
ISI	Institute for Scientific Information
MCC	Matthews Correlation Coefficient
MLP	Multilayer Perceptron
MOOCs	Massive Open Online Courses
PDF	Portable Document Format
PRC	Precision Recall
ROC	Receiver Operating Characteristic
RQ	Research Question
SL	Supervised Learning
SR	Systematic Review
ST	Stereotype Threat
STEM	Science, Technology, Engineering, and Mathematics
TP	True Positive
URL	Uniform Resource Locator
WWW	World Wide Web

CONTENTS

1 INTRODUCTION

Educational Technology (edtech), defined by (LAI, 2016) as a systematic process of applying technology to improve education, has impacted the way we learn and teach. For instance, researchers have shown technology can bolster students' engagement (CHEN; LAMBERT; GUIDRY, 2010; JR, 2002; Kearsley, G., & Shneiderman, 1998), foster collaboration (CHAN et al., 2006; BELDARRAIN, 2006), and improve learning retention (BAXTER, 2012; SIMPSON, 2003). In addition, edtechs can support teachers in organizational and technical activities (BATES; POOLE, 2003), and assist them in creating and delivering new content (MAHINI; FORUSHAN; HAGHANI, 2012). However, despite these benefits, researchers have also highlighted negative aspects presented in edtechs, in particular, the presence of gender stereotype threat.

Gender differences in educational technologies and its consequences have been studied by researchers over a couple of years. Most studies show evidence that students under stereotype threat tend to perform low (ARONSON et al., 1999a; BEILOCK; RYDELL; MCCONNELL, 2007; CHRISTY; FOX, 2014). A recent literature review (PENNINGTON et al., 2016) on psychological mediators of stereotype threat and low performance classified their psychological consequences in three main categories, Affective, Cognitive and Motivational Mechanisms. A couple of authors highlight affective consequences for individuals who belong to negatively stereotyped groups. For instance, threatened participants became more depressive in the presence of emotional regulations, which had accounted to control their expression of anxiety and impacted their performance (STEELE, 1997). Another study (WEHRENS et al., 2010) highlighted evidence of negative responses in individuals' reading comprehension. Stereotype threat also affects students' performance in math problems that rely heavily on working memory rather than when the solutions are retrieved directly from the long-term memory(BEILOCK; RYDELL; MCCONNELL, 2007). Regarding motivational mechanisms, psychological disengagement was shown (SCHMADER; MAJOR; GRAMZOW, 2001) as a result of ethnic stereotype. Increase in women self-handicapping (KELLER, 2002a) and domain dis-identification(WOODCOCK et al., 2012a) were also reported.

With the aim to tackle bias issues on Web settings, several studies were already conducted. (MOWSHOWITZ; KAWAGUCHI, 2002) defined search engine bias and a system for measuring this kind of bias while (ALKHALIFA, 2015) investigated bias on the algorithm used by Google and other search engines and found the adjacency matrix used as a basis for Page Rank may have biased spaces that need to be taken into consideration. (RIQUELME; KEGENG, 2004) and (VAUGHAN; THELWALL, 2004) found evidence for bias in different business websites which may affect customers decisions. (BOZDAG, 2013) analyzed filtering processes in content personalization and found that both human and technical biases are present in today's emergent gatekeepers.

Another study (ORDUÑA-MALEA; Luis Ortega; F. Aguillo, 2014) investigated whether both web file types and language influence the web visibility of European universities and the results show that Spanish and English correlate most with web visibility. The results presented by (WHITE; HASSAN, 2014) and (WHITE; HORVITZ, 2015) show bias in web search and retrieval in online health search which affects the beliefs about the efficacy of medical interventions when exposing a person to information on search results. Despite these advances when investigating web bias, most studies do not consider aspects that affect users' psychological mechanisms like stereotype threat.

Since the content found on the Web may vary from page to page, identifying if a page is stereotyped is a complex task due to many variables involved. To address that issue, we first gathered two sets of web pages containing gender bias, i.e., one set containing female stereotype and the other one with the male stereotype. Then, we used supervised learning to create a predictive model based on the two sets of data collected. Given the subjectivity of stereotype threat makes it hard to identify which aspects activate it. Therefore, to address the subjectivity problem, we developed an approach based on situational cues [1]. (STEELE; SPENCER; ARONSON, 2002) define stereotype threat as a situational threat that arises from situational cues signaling that a negative stereotype about one's social identities is relevant as an interpretation for one's behavior in the setting. Given that, we considered biased visual arrangements like textual content and color scheme as signals of the negative stereotype.

1.1 Goals

The general goal of this work is to create and validate an approach to identify gender bias on online educational settings. In addition, this project has the following specific objectives:

- To highlight the types of edtech that have evidence of gender stereotypes
- To find in the literature the negative consequences of gender stereotypes in edtech
- To highlight the current methodological approaches to spot gender stereotype
- To summarize the limitations of current findings regarding the impact of gender stereotypes
- To apply different supervised learning techniques to generate predictive models of gender stereotypes
- To conduct a case study in top-ranked universities' web pages and provide a report regarding the presence of gender stereotypes on them.

[1] Situational cues are contextual cues in the environment (e.g., its organization, features, and physical characteristics) that signal a person an action or event may occur.

1.2 Expected Contributions

This study contributes to a multidisciplinary extent that includes Artificial Intelligence in Education and Social Psychology. Since we are presenting an approach to predict gender bias, we hope software engineers and stakeholders can use this method in their field to improve the content of online educational applications. We also expect the operationalization of variables and the systematic review related to the theory of stereotype threat can strengthen the works already conducted in social psychology and encourage researchers and advocates of educational technology to make themselves more aware of the importance of gender stereotypes in such contexts. In addition, we hope the report on top-ranked universities presented in this work can give us an insight and make us reflect on gender bias in renowned universities around the world and encourage new research in this field. Overall, we hope to benefit stigmatized groups of individuals and contribute to creating a Web more inclusive and fair.

1.3 Thesis Outline

This thesis is organized as follows. Next chapter presents a literature search about the knowledge in which this work is based, i.e., online educational technologies, stereotype threat, and supervised learning. Chapter 3 presents a literature review performed to validate the problem and includes our research questions and protocol. We also include a subsection highlighting works that are related to this study. Chapter 4 details our proposal that consists of the development of an approach to detect gender bias in edutech and is divided into attribute selection, i.e., color and text features we used; pre-processing phase and its implementation; then, classifiers and the case study. Finally, we present the results, discuss the findings, highlight possible future works, and present our conclusion(Chapter 6).

2 BACKGROUND

2.1 Educational Technologies

According to (HSU; HUNG; CHING, 2013), edtech has evolved in multiple dimensions: as a way of dealing with learning processes (ELY, 1963), as a conceptual framework (DAVIES; SCHWEN, 1972), as theory and practice (SEELS; RICHEY, 2012), and in terms of ethical practices of dealing with technological processes and resources (RICHEY; SILBER; ELY, 2008). Here, we adopt the definition of educational technology proposed by (ROBINSON; MOLENDA; REZABEK, 2013) which is "the study and ethical practice of facilitating learning and improving performance by creating, using, and managing appropriate technological processes and resources."

There are several studies providing evidence of the positive aspects of edtech. These include the possibilities of immediate feedback, frequent testing, improving access to education and interactions among teachers and students, and enhancing students' engagement. For instance, (SCHEELER; LEE, 2002) examined the effects of immediate corrective feedback on a teaching behavior and revealed it was effective in increasing pre-service teacher course completion. Meanwhile, (PENNEBAKER; GOSLING; FERRELL, 2013) used a computer-based system with daily online testing. The findings suggested that "frequent consequential quizzing should be used routinely in large lecture courses to improve performance in class." (PINETEH, 2012), on the other hand, investigated the use of virtual interactions in a communication class and found that computer-assisted exercises improve teaching and learning of communication concepts. Other studies have also demonstrated that educational technology can be used to enhance student engagement in their learning activities (CHEN; LAMBERT; GUIDRY, 2010; SHERER; SHEA, 2011; CARLE; JAFFEE; MILLER, 2009)

On the other hand, researchers have also criticized aspects of educational technologies. For instance, (BRANSFORD; BROWN; COCKING, 2000) argue that the misuse of technology can hinder rather than enhance learning. In addition, the use of digital devices and social media may distract students during the learning process (DOUGLAS; ANGEL; BETHANY, 2012; FEWKES; MCCABE, 2012). Another important concern is the unintended incorporation of human biases into edtech (HUANG; HSU; KU, 2012; KOPCHA; SULLIVAN, 2007; ARNOTT, 2006), to which we now turn. However, those limitations may be related to stereotype threat. In fact, (APPEL; KRONBERGER; ARONSON, 2011) summarize evidence showing that negative stereotypes in different media impacts cognition and educational achievement of students of negatively portrayed groups.

Recent studies have pointed the impact of stereotype threat in educational technologies. For instance, (ALBUQUERQUE et al., 2017) and (CHRISTY; FOX, 2014) investigated how gender stereotypes affect students' anxiety and performance respectively in gamified educational settings. Other studies had also made an effort to understand the effects of ST in several settings, but before presenting their findings, we will highlight the basic concepts of stereotype threat. Next section, presents the definition of ST and also shows a framework to describe it.

2.2 Stereotype threat

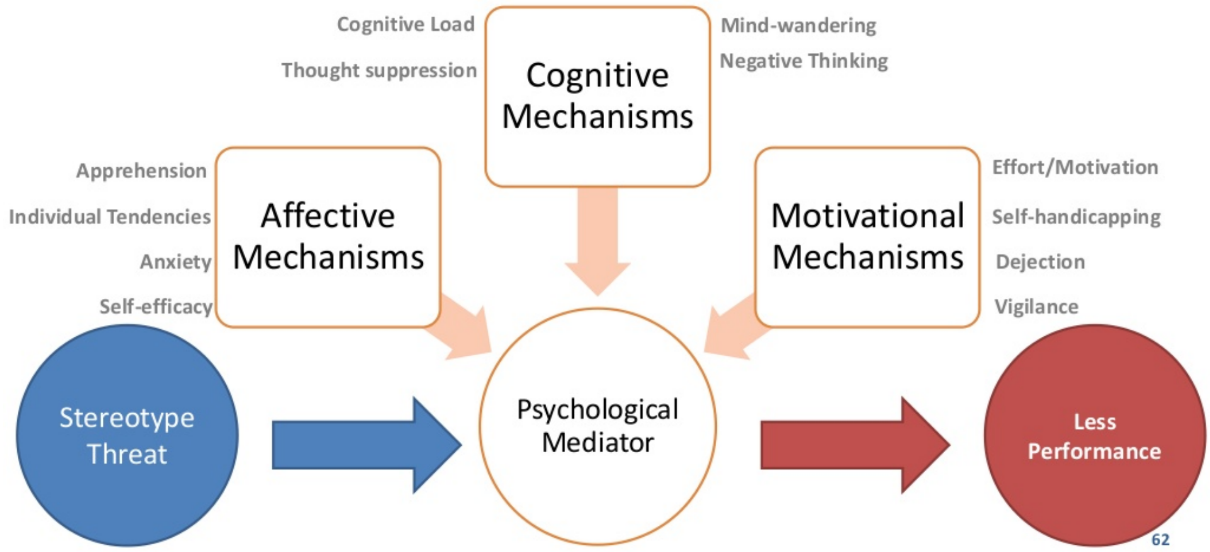
Stereotype threat is a generalized belief about a particular group of people. According to (SPENCER; STEELE; QUINN, 1999), it is a type of situational threat that arises from situational cues. Those cues indicate that a negative stereotype is relevant as an interpretation for an individual's behavior in the setting (STEELE; SPENCER; ARONSON, 2004). As an example, we may consider a civil engineering class where a teacher is showing a video of male engineers working in a construction project; in this example, the video works as a cue for female students signaling the stereotype that engineering is a field for males. (SHAPIRO; NEUBERG, 2007) conceptualized the Multi-Threat Framework in which the types of ST are the result of the interaction between two dimensions, the target and the source of stereotype threat. **Table 1** presents more detail of each resulting type of stereotype.

Table 1 – Types of Stereotype Threat according to the Multi-Threat Framework (adapted from (SHAPIRO; NEUBERG, 2007))

		Target	
		Self	**Group**
Source	**Self**	**Self-Concept Threat**. Fear that my behavior will confirm, in my own mind, that the negative stereotypes held of my group are true of me.	**Group-Concept Threat.** Fear that my behavior will confirm, in my own mind, that the negative stereotypes held of my group are true of my group
	Outgroup Members	**Own-Reputation Threat (Outgroup)**. Fear that my behavior will confirm, in the minds of outgroup members, that the negative stereotypes held of my group are true of me, and I will, therefore, be judged or treated badly by outgroup members.	**Group-Reputation Threat (Outgroup)**. Fear that my behavior will confirm, in the minds of outgroup members, that the negative stereotypes held of my group are true of my of my group and my group will, therefore, be judged or treated badly by outgroup members.
	Ingroup Member	**Own-Reputation Threat (Ingroup)**. Fear that my behavior will confirm, in the minds of ingroup members, that the negative stereotypes held of my group are true of me and I will, therefore, be judged or treated badly by ingroup members.	**Group-Reputation Threat (Ingroup)**. Fear that my behavior will confirm, in the minds of outgroup members, that the negative stereotypes held of my group are true of my group and my group will, therefore, be judged or treated badly by outgroup members.

Stereotype threat can be present in many domains, such as academic life, sport, communication, and the workplace. Most studies have investigated the effects of ST in academic environments. For instance, (CROIZET; CLAIRE, 1998) show how the performance of participants from low socioeconomic backgrounds was affected. Other authors have shown evidence of stereotypes affecting the performance of Hispanic students (GONZALES; BLANTON; WILLIAMS, 2002; SCHMADER; JOHNS, 2003), and females in Math (SPENCER; STEELE; QUINN, 1999; INZLICHT; BEN-ZEEV, 2000). (STONE, 2002) conducted two experiments to investigate the effect of ST for black and white athletes. In the first experiment, they found that white athletes performed higher when a golf activity was framed as a diagnosis of "sports intelligence". In the second experiment, black athletes scored higher when the task was framed as a diagnosis of "natural athletic ability." Meanwhile, (HIPPEL et al., 2011) showed that women changed their communication style when exposed to stereotypes regarding male leadership abilities; and in economics, (KRAY; GALINSKY; THOMPSON, 2002) investigated stereotype regeneration based on the gender gap in negotiations.

(PENNINGTON et al., 2016) conducted a systematic literature review to examine the mediators of stereotype threat and classified them into three different mechanisms, affective, cognitive, and motivational. Within the group of affective mechanisms, several studies have reported anxiety, individuation tendencies, evaluation apprehension, performance expectations, explicit stereotype endorsement, and self-efficacy as mediators. Other authors have reported the following cognitive mediators, working memory, cognitive load, thought suppression, mind-wandering, negative thinking, cognitive appraisal, and implicit stereotype endorsement. The third group includes motivation, self-handicapping, dejection, vigilance, and achievement goals as motivational mediators of stereotype threat. **Figure 1** summarizes those findings.



Source: The author

Figure 1 – Overview of Pennington's findings

In order to identify the impact of stereotype threat on students' learning, we searched the literature for the main concepts related to it. Therefore, we considered the following aspects, types of stereotype threat, its consequences, and the contexts where the studies were conducted. Below, we present the state-of-the-art regarding each of these concepts.

2.2.1 Types of stereotype threat

Many works already developed are related to stereotyping by gender. (INZLICHT; BEN-ZEEV, 2000) found that females performed lower than men in a stereotyped environment and women performance was proportionally inverse to the number of men. Other studies, for instance, (MARX; ROMAN, 2002; KELLER, 2002b; MCINTYRE; PAULSON; LORD, 2003; BEILOCK; RYDELL; MCCONNELL, 2007), considered females math performance and reported evidence showing they performed lower than males under stereotype threat.

Researchers also investigated stereotype threat regarding social groups. (ZANNA; GOETHALS; HILL, 1975) investigated social comparison among different groups of participants and found they used to compare themselves, first, to participants of the same gender and, second, to participants of the same group. (MENDES et al., 2001) studied the impact of social comparison in participants' cardiovascular responses and found the group under stereotype threat showed exacerbated reactions relative to the others. Furthermore, (STEELE; SPENCER; ARONSON, 2002) investigated individual and group stereotype and highlighted it affected negatively participants' behavior. Lately, (WEHRENS et al., 2010) found the level of empathic, constructive, and destructive responses from participants over different social groups vary significantly. Recently, (HAMARI; KOIVISTO, 2015) investigated how social influence affects participants when using gamified services and reported positive results.

Other studies also considered the effect of stereotype threat over groups with a different ethnic origin. For instance, (STEELE, 1997) reported low performance of African Americans under stereotype threat, and (WOODCOCK et al., 2012b) found that chronic stereotype led participants to domain dis-identification when considering their race and ethnicity.

2.2.2 Consequences of stereotype threat

There are many consequences of stereotype threat pointed out in the literature. Most studies show evidence that participants under stereotype threat tend to perform low (ARONSON et al., 1999a; INZLICHT; BEN-ZEEV, 2000; MARX; ROMAN, 2002; KELLER, 2002b; MCINTYRE; PAULSON; LORD, 2003; SEKAQUAPTEWA; THOMPSON, 2003; BEILOCK; RYDELL; MCCONNELL, 2007; SIEVERDING; KOCH, 2009; TAYLOR et al., 2011; CHRISTY; FOX, 2014; VERMEULEN et al., 2016). (PENNINGTON et al., 2016) performed a literature review on psychological mediators and stereotype threat and classified the psychological consequences of stereotype threat in three main categories, Affective, Cognitive and Motivational Mechanisms.

A couple of authors highlight affective consequences of stereotype threat. (STEELE, 1997) show stereotype threat made participants more depressive. (JOHNS; INZLICHT; SCHMADER, 2008) found the presence of emotional regulations among threatened participants. (WEHRENS et al., 2010) show evidence of negative responses when investigating affective mechanisms.

Other authors investigated cognitive mediators. For instance, (SCHMADER; JOHNS, 2003) investigated the impact of stereotype threat in working memory capacity. Lately, (BEILOCK; RYDELL; MCCONNELL, 2007) found that found the stereotype threat was more prone when the participants were asked to solve problems that rely heavily on working memory and can be alleviated when the problem's solutions are retrieved directly from long-term memory.

Regarding motivational mechanisms, (SCHMADER; MAJOR; GRAMZOW, 2001) investigate psychological disengagement as a result of ethnic stereotype, (KELLER, 2002b) found stereotype threat increased women self-handicapping, and (WOODCOCK et al., 2012b) reported that chronic stereotype led to domain dis-identification.

2.2.3 Contexts and domains

Many authors studied the effect of stereotype threat in the Mathematics domain (ARONSON et al., 1999b; SPENCER; STEELE; QUINN, 1999; KELLER, 2002b; MARX; ROMAN, 2002; MCINTYRE; PAULSON; LORD, 2003; IDENTITY. . . , 2004; CHRISTY; FOX, 2014). Other authors, conducted their work in other particular contexts, for instance, (CHERYAN; MELTZOFF; KIM, 2011; CHRISTY; FOX, 2014) considered the context of a virtual classroom, and (CUNDIFF et al., 2013) considered the stereotype associated to men's career aspiration in STEM domains.

2.3 Supervised Learning

Along the years, there were many efforts to make computers capable to learn from experience and automatically enhance their programs' efficiency during execution (MICHIE et al., 1994), which is a field widely known as Machine Learning. In that area of study, every instance present in a set of data can be represented by using the same set of attributes (KOTSIANTIS; ZAHARAKIS; PINTELAS, 2007) (also called *features*). Given that, when the instances have known labels that correspond to correct outputs the learning is called supervised when it does not happen the learning is called unsupervised (KOTSIANTIS; ZAHARAKIS; PINTELAS, 2007). Supervised Learning problems are commonly grouped into two categories, (i) Classification, when the expected variable is categorical [1], and (ii) Regression, when the output variable is numerical (continuous). On the other hand, Unsupervised Learning is associated with clusterization, i.e., problems where it is required to find patterns in the data. **Figure 2** shows an overview of those concepts.

2.3.1 Applications

There are many applications of both types of learning in several areas, e.g., Health (KOUROU et al., 2015; LONGSTAFF; REDDY; ESTRIN, 2010; WANG et al., 2012), Chemistry (GOH et al., 2018; GILMER et al., 2017), Business (JEYAPRIYA; SELVI, 2015; SIRIGNANO; CONT, 2018; GAI; QIU; ELNAGDY, 2016), Engineering (STRUŠNIK et al., 2016; HU; LI; YANG, 2015),

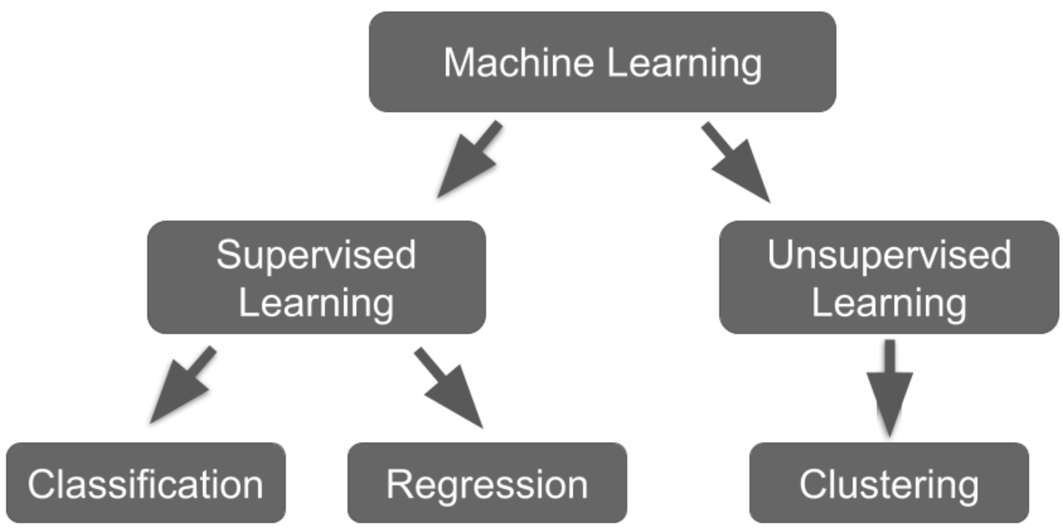
[1] In Statistics, a categorical variable is a type of variable that assumes a limited number of values (also known as *categories*) (YATES; MOORE; STARNES, 2002)

Education (LIVIERIS et al., 2018; PAPERNOT et al., 2016; LIANG; LI; ZHENG, 2016; OLIVÉ et al., 2019), and Psychology (PATEL; KHALAF; AIZENSTEIN, 2016; DENG; SCHULLER, 2012). Since the scope of this work is concentrated in supervised learning and stereotype threat in the context of education, we will stick to applications of supervised learning towards the solution of problems in Education.

2.3.1.1 Supervised Learning in Education

Several studies used supervised learning techniques to prevent students' dropouts in online courses. For instance, (LI et al., 2016) proposed a semi-supervised learning model to predict dropouts in massive open online courses (MOOCs) and their approach had a better performance than other state-of-the-art techniques. Other researchers (WAGER; WANG; LIANG, 2013) also applied a semi-supervised learning approach to create an adaptive regulator algorithm based on unlabeled data from online education settings and the results suggest their technique boosted the performance of dropout training. (LIANG; LI; ZHENG, 2016) used data from students' activities in several MOOCs to create a predictive model based on supervised classification and the results show precision close to 90%.

Other studies used supervised learning to predict students' performance. (LIVIERIS et al., 2019) used two different methods based on semi-supervised algorithms to predict final examination performance and the experiment indicated that classification accuracy can be enhanced by using more unlabeled data in the predictive models.



Source: The author

Figure 2 – Supervised Learning vs Unsupervised Learning

3 PROBLEM VALIDATION

In order to validate the problem, i.e., the negative impact of gender stereotype threat in educational technologies, we decided to conduct a systematic review. Next section, present its protocol and the steps we took during the process.

3.1 Objective and Research Questions

The objective of this review is to identify what is already known in the field about gender stereotypes in educational technologies and propose a research agenda for future works. To achieve that aim, we followed the PICOC method (PETERSEN R. FELDT; MATTSSON, 2007) with the following elements, (i) **Population**, comprised of students who used educational technologies in the past 20 years; (ii) **Intervention**, that consisted of learning through computer-based educational technologies; (iii) **Comparison** in which were gender stereotyped educational technologies; (iv) **Outcome**, consisting on the impact gender stereotype threat in students' learning; and (v) **Context**, that was set to educational technologies for academic purpose.

Next step defines our research questions based on the items above. The following questions were raised:

RQ1: Which types of educational technologies have evidence of gender stereotype threat?

RQ2: What are the negative consequences of gender stereotypes in the context of educational technologies?

RQ3: Which experimental approaches have been used to spot the threat of gender stereotypes in such contexts?

RQ4: What are the methodological limitations of the current findings?

3.2 Selection Criteria and Quality Assessment

Based on those questions we created the search string "*(student OR gender OR sex) AND (stereotype OR bias) AND (technology) AND (learning OR education)*" and searched the following sources, ACM Digital Library, Google Scholar, IEEE Digital Library, ISI Web of Science, Science Direct, Scopus, and Springer Link. Then, a set of inclusion and exclusion criteria were set (see **Table 2**). Since we intended to highlight concrete evidence of gender stereotypes in Edtech, we excluded abstracts, book chapters and review articles; in addition, only empirical studies in that context were included. Then, in order to make the review viable, we also excluded papers not in English and not available online. Articles of the domain of Edtech and not related to gender studies were also excluded, since they were out the scope of this research.

Finally, we set a time-window of 20 years and removed studies that did not validate the approach proposed in order assure quality.

Table 2 – Inclusion and exclusion criteria

Inclusion	Exclusion
Article and conference paper	Abstract, chapter, or review article
Empirical studies	Does not include validation of the approach
Gender stereotype in Edtech	Not available online
	Not empirical
	Not in EdTech domain
	Not in English
	Not related to gender studies
	Published before 1998 or after 2018

After applying the inclusion and exclusion criteria, the quality assessment checklist was applied toward the remaining articles (see Appendix A.1). Each criterion was evaluated as “Yes” (weight=1.0), “No” (weight=0.0), or “NA/NR” (weight=0.5) when a given criterion did not fit with the article or the information was not present. In addition, a cutoff score of 7.0 was defined so that articles with a score below 7.0 were eliminated.

3.3 Data extraction

Finally, a data extraction form was created to collect information according to each research question. The extracted data comprised, (i) type of educational technology; (ii) were the study was applied; (iii) impact of stereotype threat; (iv) experimental approach; and (v) limitations.

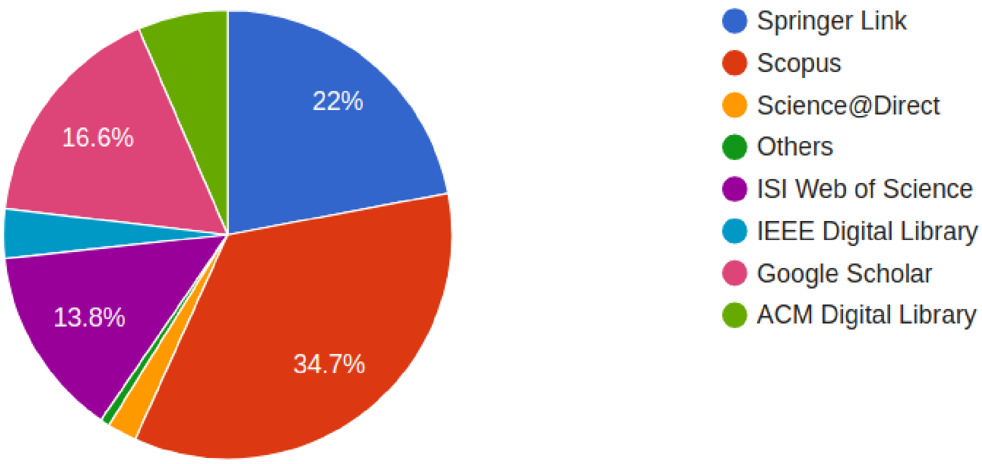
Since the protocol was defined, the reviewing process started and the results are presented in the next section. We highlighted descriptive data in the first subsection and organized the following subsections according to each research question.

3.4 Study Selection

A total of 1184 articles were initially collected (n=1184), where 8 articles came from our personal databases and 1176 from online datasets. Then, 137 duplicate entries were removed and the remaining 1047 were screened for title and abstract. After applying the exclusion criteria, 954 articles were excluded and the remaining 93 were assessed based on the inclusion criteria which resulted in 34 articles selected. Finally, the quality assessment checklist was applied and 24 articles were ultimately considered for the data extraction phase. **Figure 3** presents the proportion of studies imported by source and **Figure 4** shows the number of selected studies along the years. Notice that most papers were returned from Scopus, and the years 2012, 2015, and 2017 had more studies that attended this review’s selection criteria.

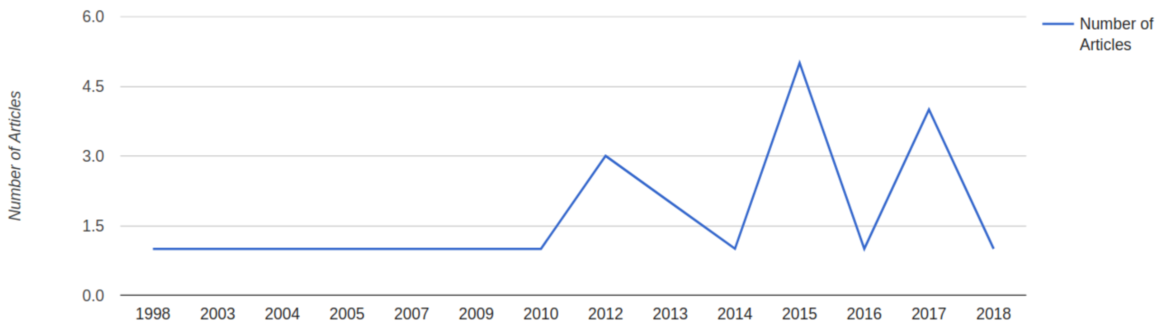
3.5 Answer to RQ1: *Which types of educational technologies have evidence of gender stereotype threat?*

The resulting articles show evidence of works in multiple types of educational technologies. For instance, online technologies, mobile and desktop applications, and others. Below, we present the evidence found on each type. Several studies have investigated gender stereotypes in **online environments** (MOSS; GUNN, 2007; SULLIVAN et al., 2015; WLADIS; CONWAY; HACHEY, 2015; JIMÉNEZ-CORTÉS; VICO-BOSCH; REBOLLO-CATALÁN, 2017; VÁZQUEZ-CANO; MENESES; GARCÍA-GARZÓN, 2017). For instance, (MOSS; GUNN, 2007) investigated gender differences in websites' design and found that men have a tendency



Source: Parsifal

Figure 3 – Imported studies per source



Source: Parsifal

Figure 4 – Selected articles per year

to prefer websites developed by men and women for websites developed by women. (SULLIVAN et al., 2015) studied gendered discuss style and showed that ideas were taken up at similar rates regardless of the gendered discourse style employed. Another study (WLADIS; CONWAY; HACHEY, 2015) investigated online courses and suggested that students in STEM may need extra support in online environments. Recently, (JIMÉNEZ-CORTÉS; VICO-BOSCH; REBOLLO-CATALÁN, 2017) studied female student in information and communication technologies and their influence on digital competence; the findings suggest women attained more advanced skills when using a wider variety of learning strategies. (VÁZQUEZ-CANO; MENESES; GARCÍA-GARZÓN, 2017) investigated gender differences in basic digital competences and the results show that men have greater perceived competence in digital cartography and online presentations while women prefer personal tutorials, have more perceived competence in corporate email and opted for personal tutors more than men.

(BROSNAN, 1998; HARTLEY; PENNEBAKER; FOX, 2003; SHELDON, 2004; SCHRODT; TURMAN, 2005; JOHNSON; GARDNER, 2009; YAU; LAI; CHENG, 2010; SHIBAZAKI; MARSHALL, 2013) investigated gender bias in **desktop applications**. For instance, (BROSNAN, 1998) studied students anxiety when using computers and the results suggested that females use males for computer-related support as a strategy to reduce anxiety. (HARTLEY; PENNEBAKER; FOX, 2003) used computer-based measures to assess students' learning styles and found no evidence to support that there would be differences between the readability of papers produced by participants with different genders. (SHELDON, 2004) studied stereotypes in preschool educational software and the findings showed more male characters than female characters in such software. (SCHRODT; TURMAN, 2005) investigated the perception of instructors' credibility based on instructional technologies and course design and found that students' perceptions of instructor's credibility changed based on sex. (JOHNSON; GARDNER, 2009) used a computer-based tutorial on car engines, manipulated a gender synthesized voice and found that females in a positive mood had a greater propensity to gender-stereotype than female in a negative one. Lately, (YAU; LAI; CHENG, 2010) used higher education technology like AutoCAD, SPSS, Compiere, Arena, and programming language (e.g., Java and Visual Basic) to assess students' confidence and the findings confirmed that males are more confident when using such tools. (SHIBAZAKI; MARSHALL, 2013) investigated gender differences in musical composition through a musical notation software package and highlighted that gender variations do exist in children's approaches to computer-based musical composition activities and that their attitudes may vary based on gender.

Two studies (ZUALKERNAN, 2015; SPIELER et al., 2018) considered **mobile applications** to analyze gender stereotypes in education. (ZUALKERNAN, 2015) looked for gender differences in technology-based numeracy interventions and found that children's gender had no significant changes in learning gain, but teachers' gender has impacted students' learning gain. (SPIELER et al., 2018) used a coding app to analyze females' performances in coding activities along two years and the evidence show that project, game-based environments, the possibility of self-expression, collaboration, and creativity can reinforce gender equity.

Other studies (CHRISTY; FOX, 2014; RICHARD; HOADLEY, 2015; ALBUQUERQUE et al., 2017; KHAN; AHMAD; MALIK, 2017) investigated the impact of stereotype threat in the context of **game-based environments**. (CHRISTY; FOX, 2014) used a virtual classroom and manipulated a leaderboard to assess the impact of stereotype threat in females performance. Another study (RICHARD; HOADLEY, 2015) considered online digital game environments to understand how supportive communities can improve resilience by mitigating stereotype threat and found that a female-supportive community can improve such settings. (ALBUQUERQUE et al., 2017) manipulated gender stereotypes in a prototype of gamified educational technology in order to check its effects in participants' anxiety. (KHAN; AHMAD; MALIK, 2017) used digital game-based learning and gamification to investigate students' engagement and found that girls outperformed boys in terms of engagement and learning outcomes.

Other studies (SCHROEDER; ADESOPE, 2015; KRÄMER et al., 2016) investigated the impact of gender and **pedagogical agents**. For instance, (SCHROEDER; ADESOPE, 2015) studied the effects of pedagogical agents on learning and the findings indicate the effects of learning with pedagogical agents may be independent of the agents' gender. On the other hand, (KRÄMER et al., 2016) investigated the effects of positive relationship interactions of a virtual agent in participants' motivation and performance in STEM and found that an agent of the opposite gender to the participants' is most successful in increasing a participant's performance.

Finally, (JONG; LU; WANG, 2010; HUANG; HSU; KU, 2012; SAUTER, 2012; HUFFMAN; WHETTEN; HUFFMAN, 2013) investigated gender bias in **educational technology in general**. (JONG; LU; WANG, 2010) proposed a model based on the acceptance of gender roles in the acceptance of information technologies and the findings indicate significant differences when comparing females in different college years. (HUANG; HSU; KU, 2012) examined the impact of confirmation bias in a computer-supported decision-making context and the results suggest that computer-mediated counter-arguments can reduce the impact of confirmation bias and lead to higher satisfaction of students. (SAUTER, 2012) tested the existence of gender stereotype threat in a Master's in Management Information Systems and the evidence suggested that positive and supportive messages have more effect on females than negative messages. (HUFFMAN; WHETTEN; HUFFMAN, 2013) investigated the influence of gender roles on technology self-efficacy and found that masculinity is the source of gender differences in technology, not the biological sex by itself.

3.6 Answer to RQ2: *What are the negative consequences of gender stereotype in the context of educational technologies?*

The studies selected in this review presented evidence of different implications of stereotype threat and gender bias. For instance, (HUFFMAN; WHETTEN; HUFFMAN, 2013), (ZUALKERNAN, 2015), and (CHRISTY; FOX, 2014) showed impacts on students' **performance**. (BROSNAN, 1998) and (ALBUQUERQUE et al., 2017) presented evidence of stereotype threat affecting **anxiety**. (SAUTER, 2012; YAU; LAI; CHENG, 2010; SHIBAZAKI; MARSHALL, 2013) indicated that **confidence** was affected. Other studies, highlighted the impact of gender bias in **resilience** (RICHARD; HOADLEY, 2015; WLADIS; CONWAY; HACHEY, 2015) and **engagement** (KHAN; AHMAD; MALIK, 2017). The findings from (JIMÉNEZ-CORTÉS; VICO-BOSCH; REBOLLO-CATALÁN, 2017; VÁZQUEZ-CANO; MENESES; GARCÍA-GARZÓN, 2017; SCHRODT; TURMAN, 2005) suggest that **perceived competence** and **self-image** (JONG; LU; WANG, 2010; SHELDON, 2004) were affected. Other authors (SULLIVAN et al., 2015; HUANG; HSU; KU, 2012; SPIELER et al., 2018; HARTLEY; PENNEBAKER; FOX, 2003) presented negative effects on students' **learning goals and styles**. (SCHROEDER; ADESOPE, 2015; KRÄMER et al., 2016; JOHNSON; GARDNER, 2009) found that stereotypes impacted **cognition**, **affection**, and **mood**. Finally, (MOSS; GUNN, 2007) showed evidence that gender differences in web-site design may lead to different users' **preferences**.

3.7 Answer to RQ3: *Which experimental approaches have been used to spot the threat of gender stereotypes in such contexts?*

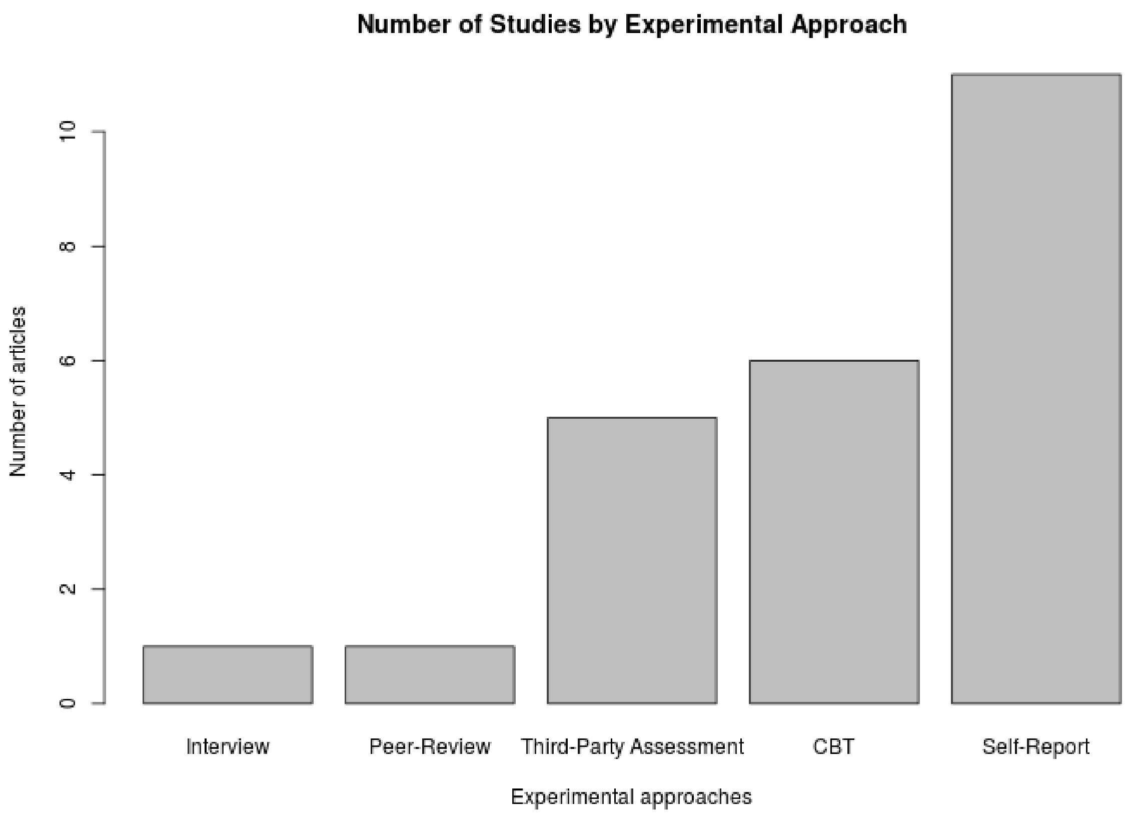
The results show that most approaches were based on a survey or self-report questionnaire. Other researchers used computer-based tests, interview, peer-review, or analyzed participants' task manually. **Figure 5** shows the method adopted for each study.

3.8 Answer to RQ4: *What are the methodological limitations of the current findings?*

Several researchers reported issues related to **sample representability** where the participants who took part in their studies did not represent a wider portion of the population. For instance, (BROSNAN, 1998) only considered students from psychology majors and (RICHARD; HOADLEY, 2015) investigated only female-supportive communities. (SHELDON, 2004) used a single software in their research and the results could not be generalized. Another study (SCHRODT; TURMAN, 2005) did not have ethnic homogeneity in their sample in which most participants were white. (KHAN; AHMAD; MALIK, 2017) performed their observations in a single classroom, with one teacher and a few students; in addition, there were too many observers and limited guidance to participants. (WLADIS; CONWAY; HACHEY, 2015) also performed their study with a single-site population and their study could not be generalized. Other researchers (SUL-

LIVAN et al., 2015) had the majority of participants formed by males since there was a lack of women and girls pursuing STEM careers. (SAUTER, 2012) chose students from the same Information System program and the results do not explain why differences seem to exist in the acceptance of females in some IT departments. (ZUALKERNAN, 2015) only considered a limited context and no significant difference was reported in the results. (SPIELER et al., 2018) experienced a high drop-out rate during the experiment and the statistical results were very poor.

Other two limitations consisted of studies who investigated either a **single domain**(HARTLEY; PENNEBAKER; FOX, 2003; CHRISTY; FOX, 2014; HUANG; HSU; KU, 2012; ALBUQUERQUE et al., 2017; KRÄMER et al., 2016) or chose a **scope too broad** (JONG; LU; WANG, 2010; HUFFMAN; WHETTEN; HUFFMAN, 2013). For instance, (HARTLEY; PENNEBAKER; FOX, 2003) investigated writing styles for classes of educational psychology only. (CHRISTY; FOX, 2014; ALBUQUERQUE et al., 2017), and (KRÄMER et al., 2016) performed experiments in a mathematics context and did not took into account other areas. Other researchers (HUANG; HSU; KU, 2012) ran their study in the financial area and used a single task consisting of two options only to test the proposed concept. On the other hand, (JONG; LU; WANG, 2010) and (HUFFMAN; WHETTEN; HUFFMAN, 2013) considered general technologies and did not consider specific settings and situations in their analyses.



Source: The author

Figure 5 – Experimental Approaches used by Imported Studies

Finally, we identified three studies with **instrument limitations**(SCHROEDER; ADESOPE, 2015; JIMÉNEZ-CORTÉS; VICO-BOSCH; REBOLLO-CATALÁN, 2017; VÁZQUEZ-CANO; MENESES; GARCÍA-GARZÓN, 2017). (SCHROEDER; ADESOPE, 2015) used text-to-speech technology to create computer-synthesized voices in their experiment which introduced distractions to participants since they could not have control of the speed and inflection of the voices. (JIMÉNEZ-CORTÉS; VICO-BOSCH; REBOLLO-CATALÁN, 2017) and (VÁZQUEZ-CANO; MENESES; GARCÍA-GARZÓN, 2017) used non-validated and non-standardized questionnaires which affects the studies' reliability.

3.9 Related Works

Next sections present an overview of the works related to this thesis, i.e., studies that tried to identify bias on the Web. Since, we found no study proposing a method to identify gender stereotype on the Web, we sought for approaches related to bias in general. For instance, we highlight studies related to bias in (i) search engines, which were conducted by (ALKHALIFA, 2015), (WHITE; HASSAN, 2014), (WHITE; HORVITZ, 2015) and (MOWSHOWITZ; KAWAGUCHI, 2002); (ii) customers' Websites, performed by (RIQUELME; KEGENG, 2004) and (VAUGHAN; THELWALL, 2004); (iii) content personalization and language developed by (BOZDAG, 2013) and (ORDUÑA-MALEA; Luis Ortega; F. Aguillo, 2014) respectively. Afterward, we highlight their limitations.

3.9.1 Bias in search engines

A few authors investigated the presence of bias in search engines. (ALKHALIFA, 2015), for example, analyzed an algorithm used by Google (PageRank) when indexing Web pages. They ran an experiment to test different formulations of the matrix used by the algorithm and the results show the matrix may contain cycles, i.e., elements capable to control the algorithm and affect the search result. They also validated the results based on mathematical proof and found the cycles persisted in the Matrix. Two other investigations (WHITE; HORVITZ, 2015; WHITE; HASSAN, 2014) also analyzed bias in search results, but in the health context. In these cases, they considered logs from search engines queries and used the results from a meta-analysis' to base their study. The findings indicate the bias present in search and retrieval in online health search affects beliefs about the efficacy of medical interventions when exposing a person to information on search results. Beyond those studies, (MOWSHOWITZ; KAWAGUCHI, 2002) presented a definition of search engine bias and describe an approach to measure it. The authors also illustrate how their method can be useful by performing statistical analysis and the results indicate that search engine bias does not depend on the search domain.

3.9.2 Bias in customers' Websites

(RIQUELME; KEGENG, 2004) and (VAUGHAN; THELWALL, 2004) pointed their investigations towards customers' Websites. (RIQUELME; KEGENG, 2004) conducted an experiment to identify different sources of bias across several business Websites. They also sought for the frequency in which such bias was observed and to what extent it is prevented. The results indicate that one-third of pages posted biased prices and the authors suggest that business may consider the implications of providing insufficient information capable to mislead customers. (VAUGHAN; THELWALL, 2004) tested national bias (US) in three different search engines regarding their relation to commercial Websites and compared to other countries like China and Taiwan. The findings suggest the results from search engines are more affected by the number of links present on each page instead of their contents by itself. The authors also highlight that bias remains an international concern on the Web.

3.9.3 Bias in content personalization and language

(BOZDAG, 2013) investigated personalization features in social Web technologies intermediate by companies like Google and Facebook. The authors studied the algorithms that filter information directed to final users and show that humans may affect filtering processes no matter if the algorithm is not only operational. They also present that personalization is related to other filtering techniques and indicate both human and technical bias in such technologies. (ORDUÑA-MALEA; Luis Ortega; F. Aguillo, 2014) investigated to what extent a set of files in different languages (English, Spanish, German, French and Italian) affected the visibility of Websites in Europe. They performed an analysis towards top-ranked universities' Websites by querying for their URLs from Google Search Engine and the results show evidence that English and Spanish pages have higher visibility. In addition, they show the correlation is stronger when considering PDF files.

3.9.4 Limitations of Related Works

Despite these advances when investigating Web bias, most studies do not consider aspects that affect users' psychological mechanisms like stereotype threat. For instance, the studies found in the literature, presented in the previous sections, are related to issues about bias in general and do not investigate gender bias specifically. In addition, those investigations do not consider psychological fundamentals of stereotype threat like situational cues. Moreover, the domain were too broad and did not take into account to develop an approach to detect gender stereotype threat in online educational technologies.

4 PROPOSAL

4.1 Gender ST detection approach

To predict gender bias, we collected a list of web pages from gender-biased websites extracted from multiple domains containing explicit stereotypes. For instance, we considered online stores specialized in products like clothes, makeup, and child's toys; in addition, we include websites of games, sports news, trends in fashion, workout programs, and courses. Afterward, we extract the text content and full-page screenshots for each website, pre-process and sanitize the data. After selecting the attributes, we apply multiple supervised learning approaches to predict female stereotype threat (ST_f), and male stereotype threat (ST_m). Then, we test the models and adjust them for better accuracy. The sections below describe each step.

4.1.1 Color Attributes

We considered colors as attributes based on previous studies that have indicated gender preferences for different types of colors. For instance, (SILVER; FERRANTE, 1995) conducted a survey asking participants to choose their preferable colors and the results indicated females had a preference for purple. On the other hand, (HALLOCK, 2003) found that blue is mostly preferable by males.

In this sense, for each screenshot, we sampled 3000 valid pixels, i.e., any pixel in which its color differs from black and white. For each one, we extract their respective colors represented as a three-dimensional array in reference to the RGB color model. Afterward, we calculate the arithmetic mean of each corresponding component so that we finish with three different values for color for each page's screenshot. **Figure 6** shows an example of how samples were collected from pages.

4.1.2 Text attributes

In order to check for stereotype threat in text content, we considered the dataset provided by (KENNISON; TROFE, 2003). In this sense, two sets of words be used, a set of words associated with females (W^{female}), and another one associated with males(W^{male}). Beyond the word list, we also use the normalized frequencies within the dataset. Afterward, we extract the text content from each set of biased web pages. After downloading the raw HTML content from each page, we applied regular expressions to remove formatting tags and unwanted characters in order to get a list of useful words. In the sequence, we exclude words with less than three characters in length and check their existence in a dictionary[1]. Finally, we compute the relative

[1] We used the American English dictionary from the Debian package Wamerican as reference

Source: The author

Figure 6 – Excerpt of a Web page where a sample of 3000 pixels were extracted

frequency of each word in the final word list (see **Equation 4.1**).

$$f_i = \frac{s_i}{max(S)} \tag{4.1}$$

Where:

- S_i is a list of words from a given web page i
- s_i is the number of occurrences of a given word in the set (S_i)
- $f_i^{(G)}$ is the relative frequency for a given word in the set $(W^{(G)})$

Finally, we compute a bias score (B) based on the number of words from each web page present in the dataset (see **Equation 4.2**).

$$f_i^{(G)} = \frac{w_i^{(G)}}{max(W^{(G)})} \tag{4.2}$$

Where:

- $W^{(G)}$ is the set of words for the stereotype G
- $w_i^{(G)}$ is the number of occurrences of a word in the set $(W^{(G)})$
- $f_i^{(G)}$ is the relative frequency for a word in the set $(W^{(G)})$

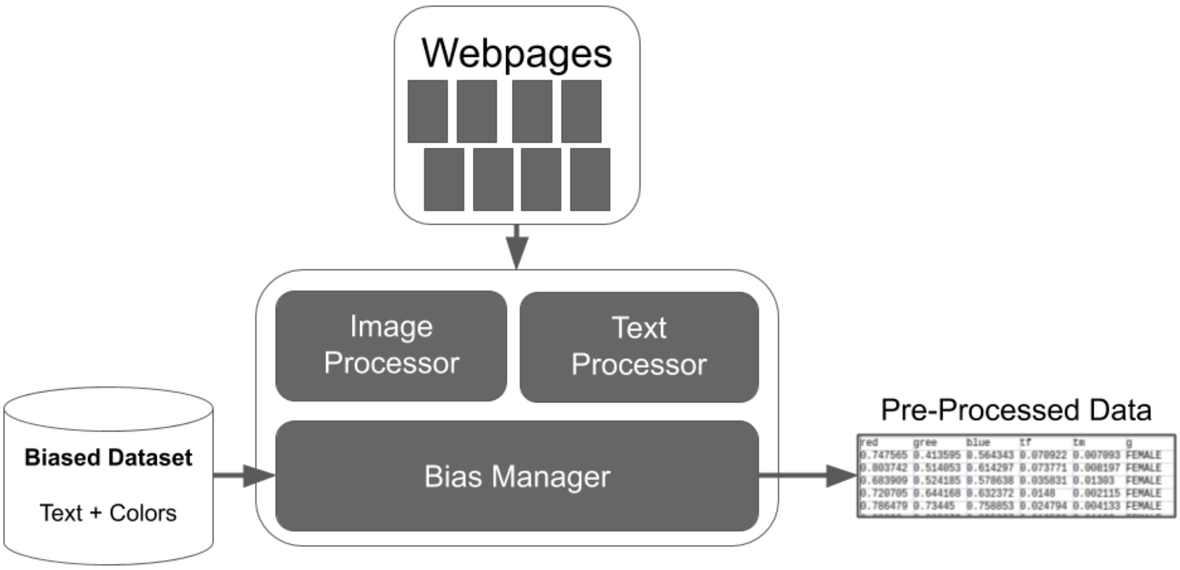
4.1.3 Pre-processing

In the pre-processing phase, we focused on two aspects, color scheme, and text content to generate a dataset containing six parameters in total. The first three parameters represent the mean of each RGB component for each page. The following two attributes represent a text score for each gender. After, the sixth attribute shows the bias - one for females and another for males. The next two sections detail how we generate each set of attributes. A.2 shows the dataset we use to generate the predictive models.

4.1.3.1 Implementation

In order to operationalize our approach, we develop a computer algorithm to implement the steps described above. The application have three main modules, (i) Image Processor; (ii) Text Processor; and (iii) Bias Manager. The Image Processor is responsible to extract screenshots from web pages and to create samples according to its colors as described in Section 4.1.1. The

Text Processor gathers text from pages, sanitize the data by removing formatting characters, calculates word frequencies, and calculates bias based on the corpus provided by (CRAWFORD et al., 2004). Finally, the Manager Module works to integrate the previous modules and to generate the final scores regarding each strategy (text and color scheme). **Figure 7** shows an overview of the architecture.



Source: The author

Figure 7 – Implementation: Pre-Processing Overview

4.1.4 Classifiers

Input Data. We apply the algorithm described in the previous subsections to generate the classifiers' input data from a sample of 100 stereotyped webpages (50 for each gender). We extract those pages from several gender-biased domains like online games, sports news, workout programs, stores, toys by gender.

Software. We use the Software Weka (University of Waikato, 2018) created by The University of Waikato to classify the data.

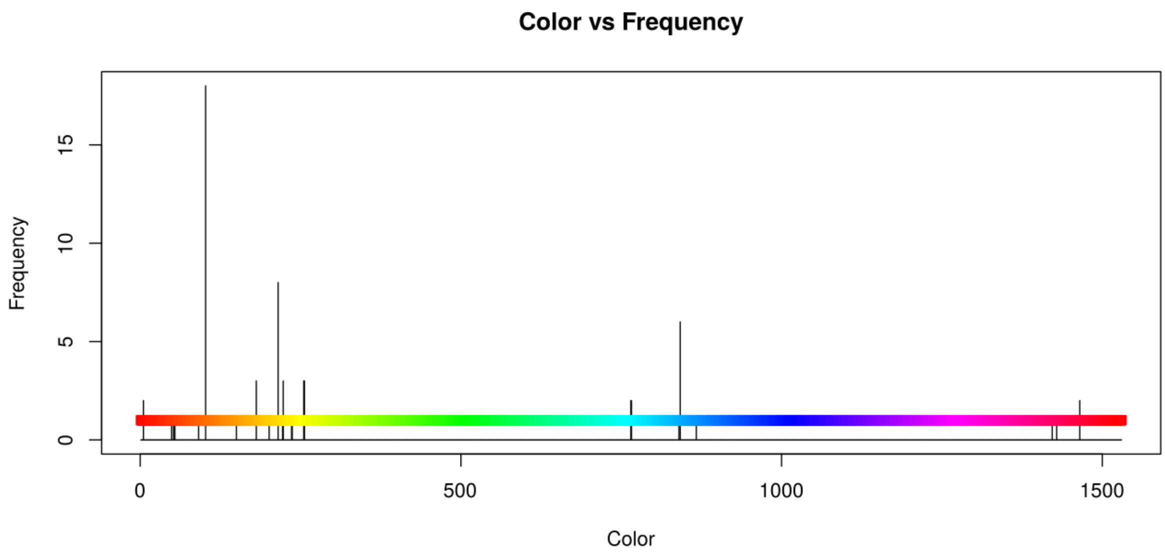
Classifiers. For this study, we generate the following Classifiers, (i) *Multilayer Perceptron algorithm (MLP)*, a neural network implementation that uses back-propagation to classify instances(ROSENBLATT, 1962; RUMELHART; HINTON; WILLIAMS, 1985); two Decision Tree algorithms to classify the training set, (ii) the *C4.5* (j48 in Weka) (QUINLAN, 2014), and (iii) the *CART (Classification and Regression Tree)* algorithm (Simple Cart in Weka) (LEO et al., 1984). In addition, we perform a *Logistic Regression* in which Weka bases its implementation on (CESSIE; HOUWELINGEN, 1992).

Setup. For all approaches, we firstly define a batch size of 100 and a precision of six decimal places. Considering the MLP, we set a learning rate of 0.3, a momentum of 0.2, a validation threshold of 20, three hidden layers, and a training time of 20,000 milliseconds. C4.5 and SimpleCart had a minimum number of observations at terminal nodes set to two, and the confidence factor for pruning be 0.25. Finally, we set a ridge of 1.0×10^{-8} and a maximum number of iterations to -1 for the Logistic Regression.

Evaluation. For each classifier, we consider a 10-fold cross-validation approach to testing the generated model. In addition, we compute the following metrics, Kappa Statistic, Mean Absolute Error, Root Mean Squared Error, Relative Absolute Error, Root Relative Squared Error, TP Rate, FP Rate, Precision, Recall, F-Measure, MCC, ROC Area, and PRC Area.

4.2 Results

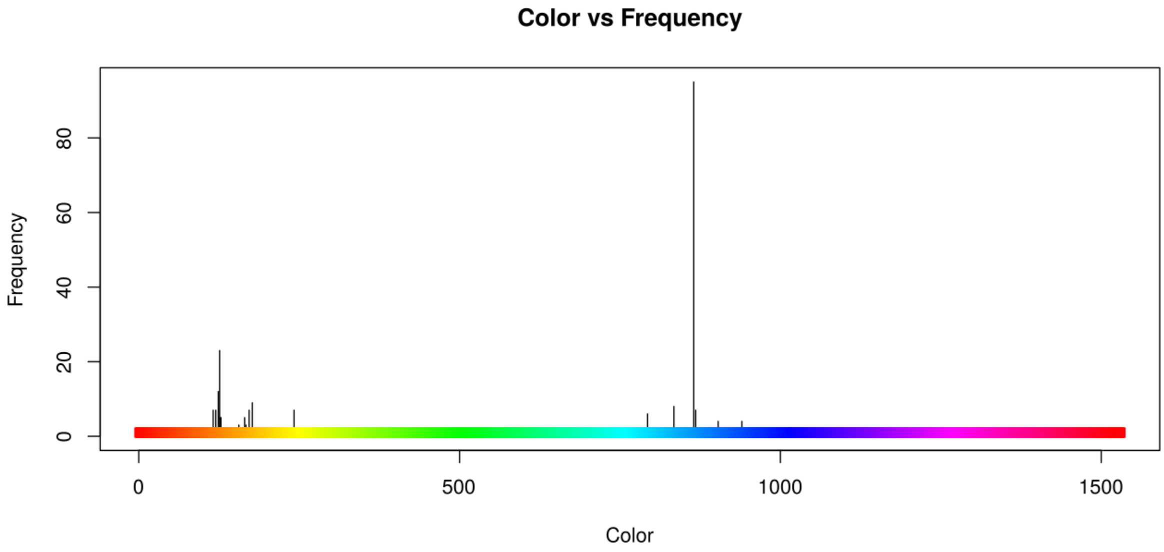
Descriptive Data For each set of web pages, we generated the correspondent color frequency according to each gender. **Figure 8** presents the corresponding frequencies for pages with female bias and **Figure 9** the frequencies for male bias. To generate those charts, we combined all colors sampled from the extracted web pages; then, we computed the corresponding frequencies and plotted them alongside a color gradient. The combination of 1530 RGB colors generated the gradient starting and ending with the color red, i.e., (255,0,0) in the RGB notation.



Source: The author

Figure 8 – Color frequencies - Female ST

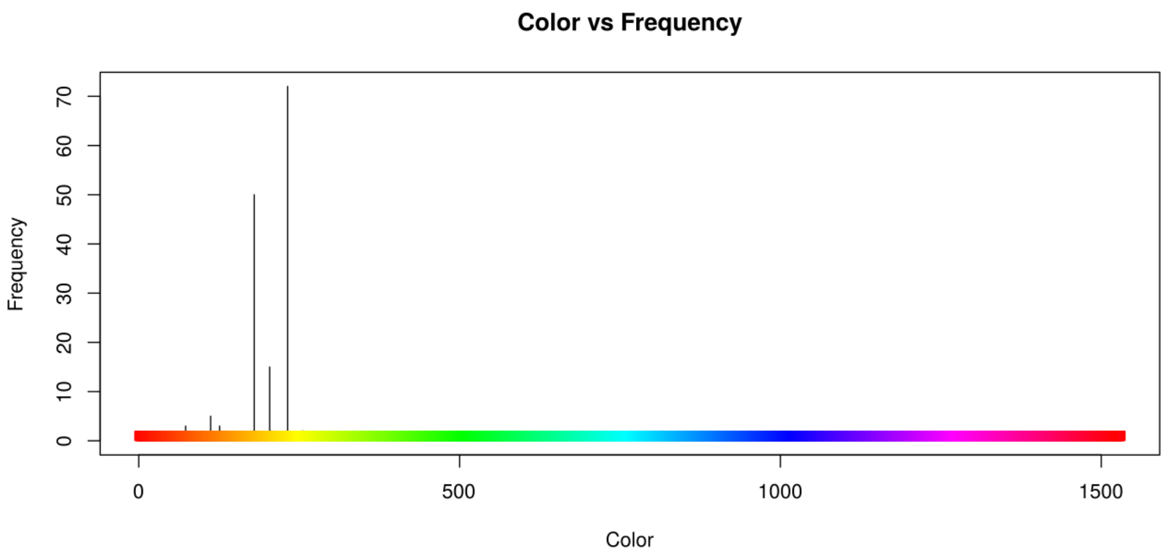
For the text content, we ranked the words and present them as word clouds. The library provided by (MUELLER et al., 2018) generated those visualizations. **Figure 11** presents top words regarding female bias and **Figure 12** male bias.



Source: The author

Figure 9 – Color frequencies - Male ST

The steps presented above were also applied to generate descriptive data for the case study. **Figure 10** shows the color frequency along the RGB gradient, and **Figure 13** depicts the most common words in a word-cloud-form.



Source: The author

Figure 10 – Color frequencies - Case Study

Source: The author

Figure 11 – Word frequencies in female-biased pages

Table 3 – Overall results for each classifier

	MLP	**C4.5**	**SimpleCart**	**Logistic**
Correctly Classified Instances	94	94	94	94
Incorrectly Classified Instances	6	6	6	6
Kappa statistic	0.88	0.88	0.88	0.88
Mean absolute error	0.08	0.06	0.08	0.07
Root mean squared error	0.24	0.25	0.24	0.25
Relative absolute error	16.32%	13.00%	16.32%	13.74%
Root relative squared error	47.71%	49.50%	47.71%	50.44%
Total Number of Instances	100	100	100	100

4.2.1 Model Evaluation

The number of correct and incorrect classified instances was the same for all classifiers (n=94 and n=6) and the Cohen’s kappa coefficient (k=0.88). The other values were very similar among the classifiers. We present a summary comparing each one in **Table 3** and the accuracy by predicted class in **Table 4**.

Source: The author

Figure 12 – Word frequencies in male-biased pages

Table 4 – Detailed Accuracy by Class. Where TP=True Positive, FP=False Positive, MCC=Matthews Correlation Coefficient, ROC=Receiver Operating Characteristic, PRC=Precision-Recall Curve.

		TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
		0.960	0.080	0.923	0.960	0.941	0.881	0.973	0.982	FEMALE
MLP		0.920	0.040	0.958	0.920	0.939	0.881	0.973	0.939	MALE
	Weighted Avg.	0.940	0.060	0.941	0.940	0.940	0.881	0.973	0.961	
		0.960	0.080	0.923	0.960	0.941	0.881	0.938	0.903	FEMALE
C4.5		0.920	0.040	0.958	0.920	0.939	0.881	0.938	0.922	MALE
	Weighted Avg.	0.940	0.060	0.941	0.940	0.940	0.881	0.938	0.912	
		0.920	0.040	0.958	0.920	0.939	0.881	0.939	0.928	FEMALE
SimpleCart		0.960	0.080	0.923	0.960	0.941	0.881	0.939	0.916	MALE
	Weighted Avg.	0.940	0.060	0.941	0.940	0.940	0.881	0.939	0.922	
		0.940	0.060	0.940	0.940	0.940	0.880	0.961	0.962	FEMALE
Logistic		0.940	0.060	0.940	0.940	0.940	0.880	0.969	0.948	MALE
	Weighted Avg.	0.940	0.060	0.940	0.940	0.940	0.880	0.965	0.955	

4.2.2 Case study

We applied the method above to top-ranked universities' web pages (n=100[2]) based on the ranking provided by (QS Quacquarelli Symonds Limited, 2018). First, we extracted screenshots and text content from each home page and processed the data following the steps

[2] After the pre-processing stage; we had n=95. Five universities' pages were excluded because we could not get a screenshot from them

Source: The author

Figure 13 – Word frequencies in universities' pages

in the previous sections. For all universities, the English version of their website was used[3]. Afterward, we applied the predictive models generated above to estimate the presence of gender stereotype in each one. Table 6 presents the list of academies and the next section shows the results.

By applying each model to the dataset from the universities' web pages, we found that most classifiers predicted male bias. For instance, the MLP classified four instances as female-biased (prediction=0.9 in two instances, and prediction=0.8 and prediction=0.6 in the other two) and 91 instances as male-biased (prediction=1.0 in 89 cases, prediction=0.9 in two cases, prediction=0.7 in two cases, and prediction=0.5 in two cases); the C4.5 algorithm outputted male bias in 84 web pages (prediction=1.0) and female bias in 11 pages (prediction=0.8); the SimpleCart approach classified all instances as male-biased (prediction=0.9); finally, the logistic regression predicted male bias in 93 cases (prediction=1.0 in 92 entries, and prediction=0.6 in one case) and female bias in two pages (prediction=1.0).

[3] Some universities in the ranking are located in countries where English is not the native language. But, all institutions considered in this study have an English version of their pages

5 DISCUSSION AND FUTURE WORKS

5.1 Discussion

In the female-biased web pages, we noticed a larger variance when compared to the spectrum of male-biased samples. Those findings are in accordance with previous studies which highlight that females have a larger color vocabulary when compared to males (NOWACZYK, 1982; SIMPSON; TARRANT, 1991). We also notice the predominance of reddish colors in the female-biased pages while in the male-biased ones the most common color was blue. Regarding the text content, we noticed some words appeared in both sets, but in different frequencies. For instance, the words *product*, *order*, and *look* were more frequent within female-biased pages while *game*, *right*, *time*, and *contact* in the male-biased ones.

After generating the models, we noticed a common similarity among their precision. In particular, the results for the MLP and SimpleCart were the same. In addition, the C4.5 and the logistic regression had similar evaluation outcomes. Furthermore, when analyzing the accuracy by each gender, we notice the MLP and C4.5 were more precise for estimating male bias, while the SimpleCart was more accurate to predict female-biased pages. Finally, the logistic regression estimated both biases with the same precision.

When applying the classifiers to top-ranked universities' homepages, we found evidence suggesting they have more male bias in their content. Those findings are aligned with the gender gap highlighted in other studies (BARBEZAT; HUGHES, 2005; GROUP, 2014; HAUSMANN, 2009; JAGSI et al., 2006). In addition, the most frequent words are related to academia (e.g., *program*, *course*, *graduate*, and *science*). We may assume this characteristic is a consequence for choosing the academic domain, i.e., universities' web pages.

5.1.1 Limitations

There are a few limitations related to the presented approach, which includes dynamicity of Web pages, automatization in portals where authentication is needed, difficult to deal with inconsistency web content, and limited number of prime variables/features. Notice that most of those limitations are related to the extraction of content from the web. More details about each one is presented below.

Content dynamicity. In web sites where its contents are updated frequently, it is hard to make any conclusions regarding the presence (or absence) of gender stereotypes. That happens because the approach proposed in this work makes its stereotype prediction based on screenshots and text content extracted in a specific instant. Therefore, what is inferred from today's content may not be applied to tomorrow's content. For instance, we may consider feed-like portals in which are updated all the time.

Automatization barriers. Another limitation is related to the automatization process of this approach. Several web sites has adopted the use of mechanisms (aka "anti-bot protection") to detect when its content is being retrieved by a bot instead of a human. Captcha, for example, is a very common technology used to deny access to bots. In addition, many portals require authentication processes in order to allow users to access their content. In such cases, the automatization of this approach may become inviable.

Inconsistent HTML Structures. Many pages are also designed without a special concert regarding best practices in web development which makes its structure insconsistent. In those case, it is needed to adapt the algorithm to each scenarios and develop a specific extraction approach to each situations.

Limited number of prime variables. Finally, we highlight a limitation related to the number of features used to generate the classification models. Given the time and effort constraints to create this approach, only two prime variables categories could be included in the classification models, i.e., color attributes and text content. Despite the good precision when predicting gender stereotypes, the addition of further variables could make the results even better.

5.2 Future Works

Based on the limitations presented above, we realise this work may be improved in future researches by considering some important factors. For instance, the automatization of this approach may include the retrieval of web pages in different time periods in order to mitigate the content dynamicity problem. In addition, other advanced techniques of artificial intelligence, like natural language processing, may also be handful when dealing with text contents. Furthermore, other aspect that can be enhanced is the consideration of multimedia elements like audio, videos and photos from pages so that the prime number of variables can be increased and the classifiers generate an even better result.

Beyond those aspects, the evidence presented in the literature review highlights several aspects of gender stereotype threat in educational technologies, and suggest future research about the following aspects, (i) dialogue with other research areas, (ii) types of educational technologies, (iii) system interactions, and (iv) source and target of stereotype.

Dialogue with other areas. In order to achieve a deeper understanding of the implications of ST, it is important to account for evidence provided by researchers in different fields of study like Social Psychology and Education. Since the stereotype threat by itself is a sub-field of Social Psychology, it is critical to understand findings regarding its psychological mechanisms and effects. In addition, based on the fact that edtech is concerned with students' learning, it is also critical to base edtech research and development in a solid base of knowledge, skills, beliefs, values, and habits of students. Finally, it is also important to be aware of developments in approaches to teaching and learning, particularly from the learning sciences.

Types of edtech. The type of edtech being developed also plays an important role in understanding and mitigating or preventing the effects of stereotype threat. For instance, Computer-Supported Collaborative Learning (CSCL) raises the issue of social interactions. In other words, in-group and out-group stereotypes (such as gender and race stereotypes) may be improved in such settings. On the other hand, the misuse of game elements in edtech can also trigger stereotype threat. For example, the presence of competition may lead to stereotypes regarding the intellectual ability of users (see (CHRISTY; FOX, 2014)). Moreover, any type of stereotype in Massive Open Online Courses (MOOCs) may by definition affect large numbers of students.

System Interactions. We also need to understand the way in which students are going to interact with the technology - in particular in terms of individual interactions and/or group collaboration. Individual interactions may mitigate the presence of self-concept threat and own-reputation threat since the source of such stereotype is either Ingroup or Outgroup (SHAPIRO; NEUBERG, 2007). On the other hand, group interactions may facilitate any of the six types of threats presented in the Multi-Threat Framework.

Source and target audience. Finally, we highlight the need to understand the source of the stereotype threat and the target (the individual or individuals under threat). Researchers have shown that many aspects of education (e.g., the method of evaluation, type of feedback, and type of interaction) can benefit the learning process when they are personalized for each individual(HOLMES et al., 2018) and avoid ST. In other words, in order to build novel educational technologies free of ST, we need to avoid “one-size-fits-all” approaches and tailor techniques that adapt to the specificity of the source and the target of the threat.

6 CONCLUSION

In this paper, we have highlighted the implications of stereotype threat for educational technologies and presented a supervised learning approach to detect gender bias in online educational settings. First, the literature revealed the fundamentals of educational technologies and stereotype threat. Then, we presented current evidence of stereotypes in educational technology through a systematic review of the literature. Regarding the supervised learning approach, we tested the method with multiple machine learning classifiers and found a precision higher than 90%. In addition, we applied the approach to top-ranked universities and the results suggest they have the male stereotype. We expect the results can signal such universities' body about the content being published in their pages. As future work, we expect to run the approach in a time-span to avoid the pitfalls of dynamic contents. In addition, we intend to improve the precision of the algorithm by adding natural language processing techniques and increasing the sample used to train the models. Moreover, we hope to expand this study to investigate a larger number of online pages, principally the ones in the domain of education. A research agenda was also presented to discuss the key aspects of gender stereotypes, comprising dialogue with other research areas, types of educational technologies, system interactions, and source and target of stereotype. Overall, this work has highlighted the importance of considering gender stereotypes in the context of educational technologies and has presented and validated an approach to detect such threats. Therefore, we hope those results can affect positively the way technology is designed in order to promote the quality and effectiveness of learning.

BIBLIOGRAPHY

ALBUQUERQUE, J. et al. Does gender stereotype threat in gamified educational environments cause anxiety? an experimental study. *Computers & Education*, Elsevier, v. 115, p. 161–170, 2017. Citado 4 vezes nas páginas 27, 37, 38, and 39.

ALKHALIFA, E. M. H. Investigating Bias in The Page Ranking Approach. In: *2015 International Conference on Information and Communication Technology Research (Ictrc)*. [S.l.: s.n.], 2015. p. 294–297. ISBN 978-1-4799-8966-9. Citado 2 vezes nas páginas 23 and 40.

APPEL, M.; KRONBERGER, N.; ARONSON, J. Stereotype threat impairs ability building: Effects on test preparation among women in science and technology. *European Journal of Social Psychology*, Wiley Online Library, v. 41, n. 7, p. 904–913, 2011. Citado na página 27.

ARNOTT, D. Cognitive biases and decision support systems development: a design science approach - Arnott - 2005 - Information Systems Journal - Wiley Online Library. *Information Systems Journal*, Wiley Online Library, v. 16, n. 1, p. 55–78, 2006. Disponível em: <http://onlinelibrary.wiley.com/doi/10.1111/j.1365-2575.2006.00208.x/pdf>. Citado na página 27.

ARONSON, J. et al. When white men can't do math: Necessary and sufficient factors in stereotype threat. *Journal of experimental social psychology*, Elsevier, v. 35, n. 1, p. 29–46, 1999. Citado 2 vezes nas páginas 23 and 30.

ARONSON, J. et al. When White Men Can't Do Math: Necessary and Sufficient Factors in Stereotype Threat. *Journal of Experimental Social Psychology*, 1999. ISSN 00221031. Citado na página 31.

BARBEZAT, D. A.; HUGHES, J. W. Salary structure effects and the gender pay gap in academia. *Research in Higher Education*, Springer, v. 46, n. 6, p. 621–640, 2005. Citado na página 53.

BATES, A. W.; POOLE, G. *Effective Teaching with Technology in Higher Education: Foundations for Success*. [S.l.]: ERIC, 2003. Citado na página 23.

BAXTER, J. A. Who am i and what keeps me going? profiling the distance learning student in higher education. *The International Review of Research in Open and Distributed Learning*, v. 13, n. 4, p. 107–129, 2012. Citado na página 23.

BEILOCK, S. L.; RYDELL, R. J.; MCCONNELL, A. R. Stereotype threat and working memory: mechanisms, alleviation, and spillover. *Journal of Experimental Psychology: General*, American Psychological Association, v. 136, n. 2, p. 256, 2007. Citado 2 vezes nas páginas 23 and 30.

BELDARRAIN, Y. Distance education trends: Integrating new technologies to foster student interaction and collaboration. *Distance education*, Taylor & Francis, v. 27, n. 2, p. 139–153, 2006. Citado na página 23.

BOZDAG, E. Bias in algorithmic filtering and personalization. *Ethics and Information Technology*, v. 15, n. 3, p. 209–227, sep 2013. ISSN 13881957. Citado 3 vezes nas páginas 23, 40, and 41.

BRANSFORD, J.; BROWN, A.; COCKING, R. R. Technology to support learning. *How people learn: Brain, mind, experience*, National Academy Press Washington, DC, p. 206–230, 2000. Citado na página 27.

BROSNAN, M. J. The impact of psychological gender, gender-related perceptions, significant others, and the introducer of technology upon computer anxiety in students. *Journal of educational computing research*, SAGE Publications Sage CA: Los Angeles, CA, v. 18, n. 1, p. 63–78, 1998. Citado 2 vezes nas páginas 36 and 38.

CARLE, A. C.; JAFFEE, D.; MILLER, D. Engaging college science students and changing academic achievement with technology: A quasi-experimental preliminary investigation. *Computers & Education*, Elsevier, v. 52, n. 2, p. 376–380, 2009. Citado na página 27.

CESSIE, S. L.; HOUWELINGEN, J. C. V. Ridge estimators in logistic regression. *Applied statistics*, JSTOR, p. 191–201, 1992. Citado na página 46.

CHAN, T.-W. et al. One-to-one technology-enhanced learning: An opportunity for global research collaboration. *Research and Practice in Technology Enhanced Learning*, World Scientific, v. 1, n. 01, p. 3–29, 2006. Citado na página 23.

CHEN, P. S. D.; LAMBERT, A. D.; GUIDRY, K. R. Engaging online learners: The impact of Web-based learning technology on college student engagement. *Computers and Education*, v. 54, n. 4, p. 1222–1232, 2010. ISSN 03601315. Citado 2 vezes nas páginas 23 and 27.

CHERYAN, S.; MELTZOFF, A. N.; KIM, S. Classrooms matter: The design of virtual classrooms influences gender disparities in computer science classes. *Computers and Education*, v. 57, n. 2, p. 1825–1835, 2011. ISSN 03601315. Citado na página 31.

CHRISTY, K. R.; FOX, J. Leaderboards in a virtual classroom: A test of stereotype threat and social comparison explanations for women's math performance. *Computers & Education*, Elsevier, v. 78, p. 66–77, 2014. Citado 8 vezes nas páginas 23, 27, 30, 31, 37, 38, 39, and 55.

CRAWFORD, J. T. et al. Champagne, beer, or coffee? a corpus of gender-related and neutral words. *Behavior Research Methods, Instruments, & Computers*, Springer, v. 36, n. 3, p. 444–458, 2004. Citado na página 46.

CROIZET, J. C.; CLAIRE, T. Extending the concept of stereotype threat to social class: The intellectual underperformance of students from low socioeconomic backgrounds. *Personality and Social Psychology Bulletin*, 1998. ISSN 01461672. Citado na página 28.

CUNDIFF, J. L. et al. Do gender-science stereotypes predict science identification and science career aspirations among undergraduate science majors? *Social Psychology of Education*, v. 16, n. 4, p. 541–554, 2013. ISSN 13812890. Citado na página 31.

DAVIES, I. K.; SCHWEN, T. M. Toward a definition of instructional development. ERIC, 1972. Citado na página 27.

DENG, J.; SCHULLER, B. Confidence measures in speech emotion recognition based on semi-supervised learning. In: *Thirteenth Annual Conference of the International Speech Communication Association*. [S.l.: s.n.], 2012. Citado na página 32.

DOUGLAS, D.; ANGEL, H.; BETHANY, W. Digital devices, distraction, and student performance: Does in-class cell phone use reduce learning? *Astronomy education review*, 2012. Citado na página 27.

ELY, D. P. The changing role of audiovisual process in education - a definition and a glossary of related terms. ERIC, 1963. Citado na página 27.

FEWKES, A. M.; MCCABE, M. Facebook: Learning tool or distraction? *Journal of Digital Learning in Teacher Education*, Taylor & Francis, v. 28, n. 3, p. 92–98, 2012. Citado na página 27.

GAI, K.; QIU, M.; ELNAGDY, S. A. Security-aware information classifications using supervised learning for cloud-based cyber risk management in financial big data. In: IEEE. *2016 IEEE 2nd International Conference on Big Data Security on Cloud (BigDataSecurity), IEEE International Conference on High Performance and Smart Computing (HPSC), and IEEE International Conference on Intelligent Data and Security (IDS)*. [S.l.], 2016. p. 197–202. Citado na página 31.

GILMER, J. et al. Neural message passing for quantum chemistry. In: JMLR. ORG. *Proceedings of the 34th International Conference on Machine Learning-Volume 70*. [S.l.], 2017. p. 1263–1272. Citado na página 31.

GOH, G. B. et al. Using rule-based labels for weak supervised learning: A chemnet for transferable chemical property prediction. In: ACM. *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. [S.l.], 2018. p. 302–310. Citado na página 31.

GONZALES, P. M.; BLANTON, H.; WILLIAMS, K. J. *The effects of stereotype threat and double-minority status on the test performance of latino women*. 2002. Citado na página 28.

GROUP, F. L. B. C. *Report: Overview of Women Status in Science*. [S.l.], 2014. Disponível em: <https://www.issuelab.org/resource/overview-of-women-status-in-science.html>. Citado na página 53.

HALLOCK, J. Colour assignment. *accessed May*, v. 1, p. 2016, 2003. Citado na página 43.

HAMARI, J.; KOIVISTO, J. Working out for likes: An empirical study on social influence in exercise gamification. *Computers in Human Behavior*, Pergamon-Elsevier Science LTD, v. 50, p. 333–347, 2015. ISSN 0747-5632. Citado na página 30.

HARTLEY, J.; PENNEBAKER, J.; FOX, C. Using new technology to assess the academic writing styles of male and female pairs and individuals. *Journal of Technical Writing*, v. 33, n. 3, p. 243–261, 2003. ISSN 0047-2816. Citado 3 vezes nas páginas 36, 38, and 39.

HAUSMANN, R. The global gender gap report 2009. In: WORLD ECONOMIC FORUM. [S.l.], 2009. Citado na página 53.

HIPPEL, C. von et al. Stereotype threat and female communication styles. *Personality and Social Psychology Bulletin*, 2011. ISSN 01461672. Citado na página 29.

HOLMES, W. et al. *Technology-enhanced Personalised Learning: Untangling the Evidence*. Stuttgart: Robert Bosch Stiftung GmbH, 2018. Disponível em: <http://oro.open.ac.uk/56692/>. Citado na página 55.

HSU, Y. C.; HUNG, J. L.; CHING, Y. H. *Trends of educational technology research: More than a decade of international research in six SSCI-indexed refereed journals*. 2013. Citado na página 27.

HU, X.; LI, S. E.; YANG, Y. Advanced machine learning approach for lithium-ion battery state estimation in electric vehicles. *IEEE Transactions on Transportation Electrification*, IEEE, v. 2, n. 2, p. 140–149, 2015. Citado na página 31.

HUANG, H. H.; HSU, J. S. C.; KU, C. Y. Understanding the role of computer-mediated counter-argument in countering confirmation bias. *Decision Support Systems*, Elsevier, v. 53, n. 3, p. 438–447, 2012. ISSN 01679236. Citado 4 vezes nas páginas 27, 37, 38, and 39.

HUFFMAN, A. H.; WHETTEN, J.; HUFFMAN, W. H. Using technology in higher education: The influence of gender roles on technology self-efficacy. *Computers in Human Behavior*, PERGAMON-ELSEVIER SCIENCE LTD, THE BOULEVARD, LANGFORD LANE, KIDLINGTON, OXFORD OX5 1GB, ENGLAND, v. 29, n. 4, p. 1779–1786, jul 2013. ISSN 07475632. Citado 3 vezes nas páginas 37, 38, and 39.

IDENTITY bifurcation in response to stereotype threat: Women and mathematics. *Journal of Experimental Social Psychology*, v. 40, n. 2, p. 152–168, 2004. ISSN 00221031. Citado na página 31.

INZLICHT, M.; BEN-ZEEV, T. A Threatening Intellectual Environment: Why Females Are Susceptible to Experiencing Problem-Solving Deficits in the Presence of Males. *Psychological Science*, 2000. ISSN 09567976. Citado 3 vezes nas páginas 28, 29, and 30.

JAGSI, R. et al. The “gender gap” in authorship of academic medical literature—a 35-year perspective. *New England Journal of Medicine*, Mass Medical Soc, v. 355, n. 3, p. 281–287, 2006. Citado na página 53.

JEYAPRIYA, A.; SELVI, C. K. Extracting aspects and mining opinions in product reviews using supervised learning algorithm. In: IEEE. *2015 2nd International Conference on Electronics and Communication Systems (ICECS)*. [S.l.], 2015. p. 548–552. Citado na página 31.

JIMÉNEZ-CORTÉS, R.; VICO-BOSCH, A.; REBOLLO-CATALÁN, A. Female university student’s ICT learning strategies and their influence on digital competence. *International Journal of Educational Technology in Higher Education*, v. 14, n. 1, 2017. ISSN 23659440. Disponível em: <http://educationaltechnologyjournal.springeropen.com/articles/10.1186/s41239-017-0040-7>. Citado 4 vezes nas páginas 35, 36, 38, and 40.

JOHNS, M.; INZLICHT, M.; SCHMADER, T. Stereotype Threat and Executive Resource Depletion: Examining the Influence of Emotion Regulation. *Journal of Experimental Psychology: General*, 2008. ISSN 00963445. Citado na página 30.

JOHNSON, D.; GARDNER, J. Exploring mindlessness as an explanation for the media equation: A study of stereotyping in computer tutorials. *Personal and Ubiquitous Computing*, Springer-Verlag, London, UK, UK, v. 13, n. 2, p. 151–163, feb 2009. ISSN 16174909. Disponível em: <http://dx.doi.org/10.1007/s00779-007-0193-9>. Citado 2 vezes nas páginas 36 and 38.

JONG, D.; LU, Y. C.; WANG, T. S. The impact of gender issues on the acceptance of information technology. In: *ICETC 2010 - 2010 2nd International Conference on Education Technology and Computer*. IEEE, 2010. v. 2, p. V2—-177—-V2—-179. ISBN 9781424463688. Disponível em: <http://ieeexplore.ieee.org/document/5529410/>. Citado 3 vezes nas páginas 37, 38, and 39.

JR, W. D. B. Student engagement, visual learning and technology: can interactive whiteboards help? Valdosta State University College of Education, 2002. Citado na página 23.

Kearsley, G., & Shneiderman, B. A framework for technology-based teaching and learning. *Education Technology*, JSTOR, v. 1, n. 1, p. 20–37, 1998. Citado na página 23.

KELLER, J. Blatant stereotype threat and women's math performance: Self-handicapping as a strategic means to cope with obtrusive negative performance expectations. *Sex Roles*, Springer, v. 47, n. 3-4, p. 193–198, 2002. Citado na página 23.

KELLER, J. Blatant stereotype threat and women's math performance: Self-handicapping as a strategic means to cope with obtrusive negative performance expectations. *Sex Roles*, Kluwer Academic Publishers-Plenum Publishers, v. 47, n. 3-4, p. 193–198, 2002. ISSN 03600025. Disponível em: <http://link.springer.com/10.1023/A:1021003307511>. Citado 2 vezes nas páginas 30 and 31.

KENNISON, S. M.; TROFE, J. L. Comprehending pronouns: A role for word-specific gender stereotype information. *Journal of psycholinguistic research*, Springer, v. 32, n. 3, p. 355–378, 2003. Citado na página 43.

KHAN, A.; AHMAD, F. H.; MALIK, M. M. Use of digital game based learning and gamification in secondary school science: The effect on student engagement, learning and gender difference. *Education and Information Technologies*, v. 22, n. 6, p. 2767–2804, 2017. ISSN 15737608. Disponível em: <http://link.springer.com/10.1007/s10639-017-9622-1>. Citado 2 vezes nas páginas 37 and 38.

KOPCHA, T. J.; SULLIVAN, H. Self-presentation bias in surveys of teachers' educational technology practices. *Educational Technology Research and Development*, Springer, v. 55, n. 6, p. 627–646, 2007. ISSN 10421629. Citado na página 27.

KOTSIANTIS, S. B.; ZAHARAKIS, I. D.; PINTELAS, P. E. Supervised Machine Learning : A Review of Classification Techniques General Issues of Supervised Learning Algorithms. *Informatica (Ljubljana)*, v. 31, p. 501–520, 2007. ISSN 03505596. Citado na página 31.

KOUROU, K. et al. Machine learning applications in cancer prognosis and prediction. *Computational and structural biotechnology journal*, Elsevier, v. 13, p. 8–17, 2015. Citado na página 31.

KRÄMER, N. C. et al. Closing the gender gap in STEM with friendly male instructors? on the effects of rapport behavior and gender of a virtual agent in an instructional interaction. *Computers and Education*, Elsevier Ltd, v. 99, n. April, p. 1–13, 2016. ISSN 03601315. Disponível em: <http://dx.doi.org/10.1016/j.compedu.2016.04.002>. Citado 3 vezes nas páginas 37, 38, and 39.

KRAY, L. J.; GALINSKY, A. D.; THOMPSON, L. Reversing the gender gap in negotiations: An exploration of stereotype regeneration. *Organizational Behavior and Human Decision Processes*, 2002. ISSN 07495978. Citado na página 29.

LAI, H. T. The Importance of Educational Research. *Language in India*, v. 16, n. 2, p. 101–105, 2016. ISSN 00319406. Disponível em: <http://ezproxy.uow.edu.au/login?url=http://search.ebscohost.com/login.aspx?direct=true{&}db=ufh{&}AN=113575599{&}site>. Citado na página 23.

LEO, B. et al. Classification and regression trees. *Wadsworth International Group*, 1984. Citado na página 46.

LI, W. et al. Dropout prediction in MOOCs using behavior features and multi-view semi-supervised learning. In: IEEE. *Proceedings of the International Joint Conference on Neural Networks*. [S.l.], 2016. v. 2016-October, p. 3130–3137. ISBN 9781509006199. Citado na página 32.

LIANG, J.; LI, C.; ZHENG, L. Machine learning application in moocs: Dropout prediction. In: IEEE. *2016 11th International Conference on Computer Science & Education (ICCSE)*. [S.l.], 2016. p. 52–57. Citado na página 32.

LIVIERIS, I. E. et al. Predicting secondary school students' performance utilizing a semi-supervised learning approach. *Journal of Educational Computing Research*, SAGE Publications Sage CA: Los Angeles, CA, p. 0735633117752614, 2018. Citado na página 32.

LIVIERIS, I. E. et al. Predicting Secondary School Students' Performance Utilizing a Semi-supervised Learning Approach. *Journal of Educational Computing Research*, v. 57, n. 2, p. 448–470, apr 2019. ISSN 0735-6331. Disponível em: <http://journals.sagepub.com/doi/10.1177/0735633117752614>. Citado na página 32.

LONGSTAFF, B.; REDDY, S.; ESTRIN, D. Improving activity classification for health applications on mobile devices using active and semi-supervised learning. In: IEEE. *2010 4th International Conference on Pervasive Computing Technologies for Healthcare*. [S.l.], 2010. p. 1–7. Citado na página 31.

MAHINI, F.; FORUSHAN, Z. J.-A.; HAGHANI, F. The Importance of Teacher's Role in Technology-Based Education. *Procedia - Social and Behavioral Sciences*, v. 46, p. 1614–1618, 2012. ISSN 18770428. Disponível em: <http://linkinghub.elsevier.com/retrieve/pii/S1877042812014772>. Citado na página 23.

MARX, D. M.; ROMAN, J. S. Female role models: Protecting women's math test performance. *Personality and Social Psychology Bulletin*, v. 28, n. 9, p. 1183–1193, 2002. ISSN 0146-1672. Citado 2 vezes nas páginas 30 and 31.

MCINTYRE, R. B.; PAULSON, R. M.; LORD, C. G. *Alleviating women's mathematics stereotype threat through salience of group achievements*. 2003. 83–90 p. Citado 2 vezes nas páginas 30 and 31.

MENDES, W. B. et al. Challenge and threat responses during downward and upward social comparisons. *European Journal of Social Psychology*, v. 31, n. 5, p. 477–497, 2001. ISSN 00462772. Citado na página 30.

MICHIE, D. et al. Machine learning. *Neural and Statistical Classification*, Technometrics, v. 13, 1994. Citado na página 31.

MOSS, G.; GUNN, R. Gender differences in website design: Implications for education. *4th International Conference on Cybernetics and Information Technologies, Systems and Applications, CITSA 2007, Jointly with the 5th International Conference on Computing, Communications and Control Technologies, CCCT 2007*, v. 1, n. 6, p. 54–59, 2007. Disponível em: <http://www.scopus.com/inward/record.url?eid=2-s2.0-84891437247{&}partnerID=40{&}md5=b6a017895fc83b4b18e18ceaea>. Citado 2 vezes nas páginas 35 and 38.

MOWSHOWITZ, A.; KAWAGUCHI, A. Bias on the web. *Communications of the ACM*, v. 45, n. 9, p. 56–60, sep 2002. ISSN 00010782. Disponível em: <http://dl.acm.org/ft{_}gateway.cfm?id=567527{&}typ>. Citado 2 vezes nas páginas 23 and 40.

MUELLER, A. et al. *Amueller/Word*$_{C}$*loud* $: Wordcloud1.5.0.Zenodo, 2018.Disponvelem :$ <>. Citado na página 47.

NOWACZYK, R. H. Sex-related differences in the color lexicon. *Language and Speech*, Sage Publications Sage CA: Thousand Oaks, CA, v. 25, n. 3, p. 257–265, 1982. Citado na página 53.

OLIVÉ, D. M. et al. A supervised learning framework: using assessment to identify students at risk of dropping out of a mooc. *Journal of Computing in Higher Education*, Springer, p. 1–18, 2019. Citado na página 32.

ORDUÑA-MALEA, E.; Luis Ortega, J.; F. Aguillo, I. Influence of language and file type on the web visibility of top European universities. *Aslib Journal of Information Management*, v. 66, n. 1, p. 96–116, 2014. ISSN 2050-3806. Disponível em: <http://www.emeraldinsight.com/doi/10.1108/AJIM-02-2013-0018>. Citado 3 vezes nas páginas 24, 40, and 41.

PAPERNOT, N. et al. Semi-supervised knowledge transfer for deep learning from private training data. *arXiv preprint arXiv:1610.05755*, 2016. Citado na página 32.

PATEL, M. J.; KHALAF, A.; AIZENSTEIN, H. J. Studying depression using imaging and machine learning methods. *NeuroImage: Clinical*, Elsevier, v. 10, p. 115–123, 2016. Citado na página 32.

PENNEBAKER, J. W.; GOSLING, S. D.; FERRELL, J. D. Daily online testing in large classes: Boosting college performance while reducing achievement gaps. *PloS one*, Public Library of Science, v. 8, n. 11, p. e79774, 2013. Citado na página 27.

PENNINGTON, C. R. et al. Twenty years of stereotype threat research: A review of psychological mediators. *PloS one*, Public Library of Science, v. 11, n. 1, p. e0146487, 2016. Citado 3 vezes nas páginas 23, 29, and 30.

PETERSEN R. FELDT, S. M. K.; MATTSSON, M. Systematic Mapping Studies in Software Engineering. *12th International Conference on Evaluation and Assessment in Software Engineering*, v. 17, n. 1, p. 33, 2007. ISSN 02181940. Disponível em: <http://www.worldscinet.com/ijseke/17/1701/S0218194007003112.html>. Citado na página 33.

PINETEH, E. A. Using virtual interactions to enhance the teaching of communication skills to information technology students. *British Journal of Educational Technology*, Wiley Online Library, v. 43, n. 1, p. 85–96, 2012. Citado na página 27.

QS Quacquarelli Symonds Limited. *TopUniversities*. 2018. Disponível em: <https://www.topuniversities.com/university-rankings/world-university-rankings/2018/>. Citado na página 50.

QUINLAN, J. R. *C4. 5: programs for machine learning*. [S.l.]: Elsevier, 2014. Citado na página 46.

RICHARD, G. T.; HOADLEY, C. Learning Resilience in the Face of Bias: Online Gaming, Protective Communities and Interest-Driven Digital Learning. *Proceedings of the 11th International Conference on Computer Supported Collaborative Learning*, n. August, p. 451–458, 2015. Citado 2 vezes nas páginas 37 and 38.

RICHEY, R. C.; SILBER, K. H.; ELY, D. P. Reflections on the 2008 aect definitions of the field. *TechTrends*, Springer, v. 52, n. 1, p. 24–25, 2008. Citado na página 27.

RIQUELME, H.; KEGENG, W. The unintended effects of hidden assumptions: biases on the internet. *Online Information Review*, v. 28, n. 6, p. 444–453, 2004. ISSN 1468-4527. Citado 3 vezes nas páginas 23, 40, and 41.

ROBINSON, R.; MOLENDA, M.; REZABEK, L. Facilitating learning. In: *Educational Technology: A Definition with Commentary*. [S.l.: s.n.], 2013. ISBN 9780203054000. Citado na página 27.

ROSENBLATT, F. Principles of neurodynamics. Spartan Book, 1962. Citado na página 46.

RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. *Learning internal representations by error propagation*. [S.l.], 1985. Citado na página 46.

SAUTER, V. L. The absence of gender differences among students in an MIS program. *Communications of the Association for Information Systems*, v. 31, n. 1, p. 86–104, 2012. ISSN 15293181. Citado 3 vezes nas páginas 37, 38, and 39.

SCHEELER, M. C.; LEE, D. L. Using Technology to Deliver Immediate Corrective Feedback to Preservice Teachers. *Journal of Behavioral Education*, 2002. ISSN 15733513. Citado na página 27.

SCHMADER, T.; JOHNS, M. Converging Evidence that Stereotype Threat Reduces Working Memory Capacity. *Journal of Personality and Social Psychology*, 2003. ISSN 00223514. Citado 2 vezes nas páginas 28 and 30.

SCHMADER, T.; MAJOR, B.; GRAMZOW, R. H. Coping with ethnic stereotypes in the academic domain: Perceived injustice and psychological disengagement. *Journal of Social Issues*, Wiley Online Library, v. 57, n. 1, p. 93–111, 2001. Citado 2 vezes nas páginas 23 and 31.

SCHRODT, P.; TURMAN, P. D. The Impact of Instructional Technology Use, Course Design, and Sex Differences on Students' Initial Perceptions of Instructor Credibility. *Communication Quarterly*, Taylor & Francis, v. 53, n. 2, p. 177–196, 2005. ISSN 0146-3373. Disponível em: <http://www.tandfonline.com/doi/abs/10.1080/01463370500090399>. Citado 2 vezes nas páginas 36 and 38.

SCHROEDER, N. L.; ADESOPE, O. O. Impacts of pedagogical agent gender in an accessible learning environment. *Educational Technology and Society*, v. 18, n. 4, p. 401–411, 2015. ISSN 14364522. Citado 3 vezes nas páginas 37, 38, and 40.

SEELS, B. B.; RICHEY, R. C. *Instructional technology: The definition and domains of the field*. [S.l.]: IAP, 2012. Citado na página 27.

SEKAQUAPTEWA, D.; THOMPSON, M. Solo status, stereotype threat, and performance expectancies: Their effects on women's performance. *Journal of Experimental Social Psychology*, Elsevier, v. 39, n. 1, p. 68–74, 2003. Citado na página 30.

SHAPIRO, J. R.; NEUBERG, S. L. From stereotype threat to stereotype threats: Implications of a multi-threat framework for causes, moderators, mediators, consequences, and interventions. *Personality and Social Psychology Review*, 2007. ISSN 10888683. Citado 3 vezes nas páginas 17, 28, and 55.

SHELDON, J. P. Gender stereotypes in educational software for young children. *Sex Roles*, v. 51, n. 7-8, p. 433–444, 2004. ISSN 03600025. Citado 2 vezes nas páginas 36 and 38.

SHERER, P.; SHEA, T. Using online video to support student learning and engagement. *College Teaching*, Taylor & Francis, v. 59, n. 2, p. 56–59, 2011. Citado na página 27.

SHIBAZAKI, K.; MARSHALL, N. A. Gender differences in computer- and instrumental-based musical composition. *Educational Research*, v. 55, n. 4, p. 347–360, 2013. ISSN 14695847. Citado 2 vezes nas páginas 36 and 38.

SIEVERDING, M.; KOCH, S. C. (Self-)Evaluation of computer competence: How gender matters. *Computers and Education*, v. 52, n. 3, p. 696–701, 2009. ISSN 03601315. Citado na página 30.

SILVER, N. C.; FERRANTE, R. Sex differences in color preferences among an elderly sample. *Perceptual and motor skills*, v. 80, n. 3 Pt 1, p. 920–922, 1995. ISSN 0031-5125. Citado na página 43.

SIMPSON, J.; TARRANT, A. W. Sex-and age-related differences in colour vocabulary. *Language and speech*, SAGE Publications Sage UK: London, England, v. 34, n. 1, p. 57–62, 1991. Citado na página 53.

SIMPSON, O. *Student retention in online, open and distance learning*. [S.l.: s.n.], 2003. ISBN 0203416562. Citado na página 23.

SIRIGNANO, J.; CONT, R. Universal features of price formation in financial markets: perspectives from deep learning. *Available at SSRN 3141294*, 2018. Citado na página 31.

SPENCER, S. J.; STEELE, C. M.; QUINN, D. M. Stereotype Threat and Women's Math Performance. *Journal of Experimental Social Psychology*, v. 35, p. 4–28, 1999. ISSN 00221031. Disponível em: <http://www.idealibrary.com>. Citado 2 vezes nas páginas 28 and 31.

SPIELER, B. et al. Reinforcing gender equality by analysing female teenagers' performances in coding activities. In: *Proceedings of the 4th Conference on Gender & IT - GenderIT '18*. {ACM} Press, 2018. p. 209–219. ISBN 9781450353465. Disponível em: <http://dl.acm.org/citation.cfm?doid=3196839.3196871>. Citado 3 vezes nas páginas 36, 38, and 39.

STEELE, C. M. A threat in the air: How stereotypes shape intellectual identity and performance. *American psychologist*, American Psychological Association, v. 52, n. 6, p. 613, 1997. Citado 2 vezes nas páginas 23 and 30.

STEELE, C. M.; SPENCER, S. J.; ARONSON, J. Contending with group image: The psychology of stereotype and social identity threat. *Advances in experimental social psychology*, ACADEMIC PRESS INC LTD, v. 34, p. 379–440, 2002. Citado 2 vezes nas páginas 24 and 30.

STEELE, C. M.; SPENCER, S. J.; ARONSON, J. Contending with group image: The psychology of stereotype and social identity threat. In: . [S.l.: s.n.], 2004. Citado na página 28.

STONE, J. Battling doubt by avoiding practice: The effects of stereotype threat on self-handicapping in White athletes. *Personality and Social Psychology Bulletin*, 2002. ISSN 01461672. Citado na página 28.

STRUŠNIK, D. et al. Energy efficiency analysis of steam ejector and electric vacuum pump for a turbine condenser air extraction system based on supervised machine learning modelling. *Applied energy*, Elsevier, v. 173, p. 386–405, 2016. Citado na página 31.

SULLIVAN, F. R. et al. Exploring the Role of ‘Gendered’ Discourse Styles in Online Science Discussions. *International Journal of Science Education*, ROUTLEDGE JOURNALS, TAYLOR & FRANCIS LTD, 4 PARK SQUARE, MILTON PARK, ABINGDON OX14 4RN, OXFORDSHIRE, ENGLAND, v. 37, n. 3, p. 484–504, feb 2015. ISSN 14645289. Citado 4 vezes nas páginas 35, 36, 38, and 39.

TAYLOR, C. A. et al. The Hillary Clinton effect: When the same role model inspires or fails to inspire improved performance under stereotype threat. *Group Processes & Intergroup Relations*, v. 14, p. 447–459, 2011. ISSN 1368-4302. Citado na página 30.

University of Waikato. *Weka 3: Data Mining Software in Java*. 2018. Disponível em: <https://www.cs.waikato.ac.nz/ml/weka/>. Citado na página 46.

VAUGHAN, L.; THELWALL, M. Search engine coverage bias: Evidence and possible causes. *Information Processing and Management*, v. 40, n. 4, p. 693–707, jul 2004. ISSN 03064573. Citado 3 vezes nas páginas 23, 40, and 41.

VÁZQUEZ-CANO, E.; MENESES, E. L.; GARCÍA-GARZÓN, E. Differences in basic digital competences between male and female university students of Social Sciences in Spain. *International Journal of Educational Technology in Higher Education*, v. 14, n. 1, 2017. ISSN 23659440. Disponível em: <http://educationaltechnologyjournal.springeropen.com/articles/10.1186/s41239-017-0065-y>. Citado 4 vezes nas páginas 35, 36, 38, and 40.

VERMEULEN, L. et al. Playing under threat. Examining stereotype threat in female game players. *Computers in Human Behavior*, Pergamon-Elsevier Science LTD, v. 57, p. 377–387, 2016. ISSN 0747-5632. Citado na página 30.

WAGER, S.; WANG, S.; LIANG, P. Dropout Training as Adaptive Regularization. p. 1–9, 2013. Disponível em: <http://arxiv.org/abs/1307.1493>. Citado na página 32.

WANG, Z. et al. Extracting diagnoses and investigation results from unstructured text in electronic health records by semi-supervised machine learning. *PLoS One*, Public Library of Science, v. 7, n. 1, p. e30412, 2012. Citado na página 31.

WEHRENS, M. J. et al. The relationship between affective response to social comparison and academic performance in high school. *Contemporary Educational Psychology*, v. 35, n. 3, p. 203 – 214, 2010. ISSN 0361-476X. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0361476X10000044>. Citado 2 vezes nas páginas 23 and 30.

WHITE, R. W.; HASSAN, A. Content Bias in Online Health Search. *ACM Transactions on the Web*, v. 8, n. 4, p. 1–33, oct 2014. ISSN 15591131. Disponível em: <http://dl.acm.org/citation.

cfm?doid=2686863.2663355>. Citado 2 vezes nas páginas 24 and 40.

WHITE, R. W.; HORVITZ, E. Belief Dynamics and Biases in Web Search. *ACM Transactions on Information Systems*, v. 33, n. 4, p. 1–46, may 2015. ISSN 10468188. Citado 2 vezes nas páginas 24 and 40.

WLADIS, C.; CONWAY, K. M.; HACHEY, A. C. The Online STEM Classroom—Who Succeeds? An Exploration of the Impact of Ethnicity, Gender, and Non-traditional Student Characteristics in the Community College Context. *Community College Review*, SAGE Publications Sage CA: Los Angeles, CA, v. 43, n. 2, p. 142–164, 2015. ISSN 19402325. Disponível em: <http://journals.sagepub.com/doi/10.1177/0091552115571729>. Citado 3 vezes nas páginas 35, 36, and 38.

WOODCOCK, A. et al. The consequences of chronic stereotype threat: Domain disidentification and abandonment. *Journal of personality and social psychology*, American Psychological Association, v. 103, n. 4, p. 635, 2012. Citado na página 23.

WOODCOCK, A. et al. The consequences of chronic stereotype threat: Domain disidentification and abandonment. *Journal of Personality and Social Psychology*, v. 103, n. 4, p. 635–646, 2012. ISSN 1939-1315. Citado 2 vezes nas páginas 30 and 31.

YATES, D.; MOORE, D. S.; STARNES, D. S. *The practice of statistics: TI-83/89 graphing calculator enhanced*. [S.l.]: Macmillan, 2002. Citado na página 31.

YAU, H. K.; LAI, A.; CHENG, F. Gender Difference of Confidence in Using Technology for Learning.pdf. *The Journal of Technology Studies*, ERIC, v. 38, n. 2, p. 74–79, 2010. Citado 2 vezes nas páginas 36 and 38.

ZANNA, M. P.; GOETHALS, G. R.; HILL, J. F. Evaluating a sex-related ability: Social comparison with similar others and standard setters. *Journal of Experimental Social Psychology*, v. 11, n. 1, p. 86–93, 1975. ISSN 10960465. Citado na página 30.

ZUALKERNAN, I. A. Gender differences in a technology-based numeracy intervention in a developing country. In: *Proceedings of the 5th IEEE Global Humanitarian Technology Conference, GHTC 2015*. 345 E 47TH ST, NEW YORK, NY 10017 USA: IEEE, 2015. p. 414–419. ISBN 9781467365611. Citado 3 vezes nas páginas 36, 38, and 39.

A APPENDIX

A.1 Systematic Review - Quality Assessment Checklist

1. Was the study described as randomized, a randomized trial, a randomized clinical trial, or an RCT?

2. Was the method of randomization adequate (i.e., use of randomly generated assignment)?

3. Was the treatment allocation concealed (so that assignments could not be predicted)?

4. Were study participants and providers blinded to treatment group assignment?

5. Were the people assessing the outcomes blinded to the participants' group assignments?

6. Were the groups similar at baseline on important characteristics that could affect outcomes (e.g., demographics, risk factors, co-morbid conditions)?

7. Was the overall drop-out rate from the study at endpoint 20% or lower of the number allocated to treatment?

8. Was the differential drop-out rate (between treatment groups) at endpoint 15 percentage points or lower?

9. Was there high adherence to the intervention protocols for each treatment group?

10. Were other interventions avoided or similar in the groups (e.g., similar background treatments)?

11. Were outcomes assessed using valid and reliable measures, implemented consistently across all study participants?

12. Did the authors report that the sample size was sufficiently large to be able to detect a difference in the main outcome between groups with at least 80% power?

13. Were outcomes reported or subgroups analyzed prespecified (i.e., identified before analyses were conducted)?

14. Were all randomized participants analyzed in the group to which they were originally assigned, i.e., did they use an intention-to-treat analysis?

A.2 Input dataset - List of gender-biased webpages

Table 5 – Links to biased web pages used to generate the input dataset

Website	Gender
http://www.girlsgogames.com/	female
http://www.girlgames.com/	female
http://www.y8.com/categories/girls	female
https://www.sisigames.com/	female
https://www.girlg.com/	female
https://www.dariagames.com/	female
https://www.girlsplay.com/	female
http://www.games2girls.com/cookinggames.htm	female
http://www.agame.com/games/beauty-games	female
http://www.gamesgames.com/games/cooking	female
https://www.maybelline.com/makeup-tips	female
https://www.kendrascott.com/	female
https://www.beautytipsonline.com/party-makeup.htm	female
https://shop.nordstrom.com/c/jewelry	female
http://www.hm.com/us/department/LADIES	female
https://www.prettylittlething.com/clothing/view-all-clothing.html	female
https://www.shopbop.com/whats-new-clothing/br/v=1/13243.htm?all	female
https://www.dresslink.com/clothing-category-811.html	female
https://www.justfab.com/clothing	female
https://www.goodhousekeeping.com/beauty/hair/a34090/best-beautiful-hair-tips/	female
http://7beautytips.com/15-hair-tips-every-girl-should-know/	female
http://www.womenshealthandfitness.com.au/health-beauty/beauty-tips/1388-7-hair-tips	female
https://nelly.com/uk/womens-fashion/	female
https://www.goodhousekeeping.com/beauty/nails/a34645/healthy-nail-care-tips/	female
https://www.rd.com/health/beauty/	female
https://www.cosmopolitan.com/uk/beauty-hair/	female
https://www.womenshealthmag.com/beauty/a19983197/nail-care/	female
https://www.womansday.com/	female
https://www.stalkbuylove.com/women-skirts/	female
https://www.lulus.com/categories/179_42/high-heels.html	female
https://www.stevemadden.com/thumbnail/WOMENS/HEELS/pc/2163/2215.uts?mode=viewall	female
http://www.boohoo.com/womens/shoes/high-heels	female
https://www.forever21.com/us/shop/catalog/category/f21/shoes_high-heels	female
https://www.sephora.com/shop/lipstick	female

https://www.maccosmetics.com/products/13854/products/makeup/lips/lipstick	female
https://colourpop.com/collections/liquid-lipsticks	female
https://www.nyxcosmetics.com/lipstick	female
https://www.guerlain.com/int/en-int/makeup/lips/lipsticks	female
https://www.shopmissa.com/collections/lipstick	female
http://www.newlook.com/row/womens/sale/c/row-womens-sale	female
https://www.burlingtoncoatfactory.com/burlingtoncoatfactory/women-56988.aspx	female
https://www.toysrus.com/products/dolls.jsp	female
https://www.target.com/c/dolls-toys/-/N-5xt90	female
http://www.newlook.com/row/womens/sale/c/row-womens-sale	female
https://www.partycity.com/toys-girls	female
https://www.urbanoutfitters.com/womens-clothing	female
https://www.toyuniverse.com.au/collections/girls-toys	female
https://www.dancedirect.com/uk/children/girls-dance-leotards	female
http://us.blochworld.com/childrens-ballet-shoes	female
https://www.browngirlsdoballet.com/	female
http://poki.com/en/boy	male
http://www.4j.com/Boy-games	male
http://www.kibagames.com/Boys-Games	male
http://www.agame.com/games/boy-games	male
http://www.cartoonnetwork.co.uk/games	male
https://talksport.com/boxing	male
https://www.bbc.com/sport/football	male
http://www.skysports.com/football	male
http://www.espn.com/soccer/	male
https://www.onlinestores.com/construction-gear-store.html	male
https://www.roblox.com/	male
https://www.muscleandstrength.com/workouts/10-week-mass-building-program.html	male
https://www.toysrus.com/products/boys-toys.jsp	male
https://www.menkind.co.uk/	male
http://www.addictinggames.com/	male
https://www.crazygames.com/	male
https://www.thestar.com.my/sport/	male
https://www.cbssports.com/live/	male
http://www.nadaguides.com/Cars/Body-styles	male
https://www.autotrader.co.uk/content/advice/what-s-a-category-c-or-category-d-car	male
https://www.carmax.com/cars	male
https://mensgear.net/	male
https://www.mantality.co.za/stuff.html	male

https://gearmoose.com/	male
https://www.bodybuilding.com/content/beginner-chest-training-guide.html	male
https://gamefaqs.gamespot.com/	male
http://www.dedegames.com/fighting.htm	male
http://www.85play.com/shooting-games.html	male
http://www.nbcolympics.com/news/swimming-101-rules	male
https://www.chordbuddy.com/guitar-learning-system-for-everyone/how-to-play-the-guitar-for-beginners/	male
http://www.fwi.co.uk/	male
https://www.tobys.com/collections/basketball	male
https://www.autotrainingcentre.com/automotive-online-training/auto-mechanics-online-course/	male
https://www.carmechanicsmag.co.uk/	male
http://www.themechanicdoctor.com/how-to-learn-auto-mechanics-online/	male
https://fly8ma.com/	male
https://www.military.com/	male
http://www.atpworldtour.com/	male
https://www.reddit.com/r/tennis/	male
https://www.foxsports.com/soccer	male
http://ukelite.com/2016/12/01/15-random-soccer-facts-americans-dont-know/	male
http://www.ibsasport.org/sports/swimming/	male
https://www.toolsofmen.com/	male
https://www.bbc.com/sport/boxing	male
https://shop.dallascowboys.com/Cowboys-Catalog/Clearance/c/clearance	male
https://heavy.com/social/2015/06/top-50-best-cool-toys-for-boys-sale-new-2015-batman-cars-guns-tinker/	male
http://www.carters.com/carters-kid-kid-boy-toyss	male
http://www.fashionbeans.com/	male
https://www.gq.com/style	male
https://www.esquire.com/style/	male

A.3 Case study: top-ranked universities

Table 6 – Universities

Institution	Link
Massachusetts Institute of Technology	http://web.mit.edu/
Stanford University	https://www.stanford.edu/
Harvard University	https://www.harvard.edu/
California Institute of Technology	http://www.caltech.edu/
University of Cambridge	https://www.cam.ac.uk/
University of Oxford	http://www.ox.ac.uk/
University College London	https://www.ucl.ac.uk/
Imperial College London	https://www.imperial.ac.uk/
The University of Chicago	https://www.uchicago.edu/
ETH Zurich	https://www.ethz.ch/en.html
Nanyang Technological University	http://www.ntu.edu.sg/Pages/home.aspx
École Polytechnique Fédérale de Lausanne	https://www.epfl.ch/
Princeton University	https://www.princeton.edu/
Cornell University	https://www.cornell.edu/
National University of Singapore	http://www.nus.edu.sg/
Yale University	https://www.yale.edu/
Johns Hopkins University	https://www.jhu.edu/
Columbia University - New York	https://www.columbia.edu/
University of Pennsylvania	https://www.upenn.edu/
The Australian National University	http://www.anu.edu.au/
University of Michigan	https://umich.edu/
Duke University	https://www.duke.edu/
The University of Edinburgh	https://www.ed.ac.uk/
King's College London	https://www.kcl.ac.uk/
Tsinghua University	http://www.tsinghua.edu.cn/publish/thu2018en/index.html
The University of Hong Kong	https://www.hku.hk/
Berkeley University of California	https://www.berkeley.edu/
The University of Tokyo	https://www.u-tokyo.ac.jp/en/
Northwestern University	https://www.northwestern.edu/
Hong Kong University of Science and Technology	http://www.ust.hk/
University of Toronto	https://www.utoronto.ca/
McGill University	https://www.mcgill.ca/
University of California	http://www.ucla.edu/
The University of Manchester	http://www.manchester.ac.uk/

London School of Economics and Political Science	http://www.lse.ac.uk/
Kyoto University	https://www.kyoto-u.ac.jp/en/
Seoul National University	http://www.useoul.edu/
University of California - San Diego	https://ucsd.edu/
Peking University	http://english.pku.edu.cn/
Fudan University	http://www.fudan.edu.cn/en/
Korea Advanced Institute of Science and Technology	http://www.kaist.edu/html/en/index.html
University of Melbourne	https://www.unimelb.edu.au/
École Normale Supérieure	http://www.ens.fr/en
University of Bristol	http://www.bristol.ac.uk/
University of New South Wales Sydney	https://www.unsw.edu.au/
he Chinese University of Hong Kong	http://www.cuhk.edu.hk/english/index.html
Carnegie Mellon University	https://www.cmu.edu/
The University of Queensland	https://www.uq.edu.au/
City University of Hong Kong	http://www.cityu.edu.hk/
The University of Sydney	https://sydney.edu.au/home.html
The University of British Columbia	https://www.ubc.ca/
New York University	https://www.nyu.edu/
Brown University	https://www.brown.edu/
Delft University of Technology	https://www.tudelft.nl/en/
University of Wisconsin - Madison	https://www.wisc.edu/
Tokyo Institute of Technology	https://www.titech.ac.jp/english/
Delft University of Technology	https://warwick.ac.uk/
University of Amsterdam	http://www.uva.nl/en/home
École Polytechnique - Paris	https://www.polytechnique.edu/en
Monash University	https://www.monash.edu/
University of Washington	https://www.washington.edu/
Shanghai Jiao Tong University	http://en.sjtu.edu.cn/
Osaka University	http://www.osaka-u.ac.jp/en
Technical University of Munich	https://www.tum.de/en/
University of Glasgow	http://www.gla.ac.uk/
Ludwig-Maximilians-Universität München	https://www.en.uni-muenchen.de/
The University of Texas at Austin	https://www.utexas.edu/
Heidelberg University	https://www.uni-heidelberg.de/index_e.html
University of Illinois	http://illinois.edu/
Georgia Institute of Technology	http://www.gatech.edu/
Pohang University of Science and Technology	http://www.postech.ac.kr/eng/

Katholieke Universiteit Leuven	https://www.kuleuven.be/english/
University of Zurich	http://www.uzh.ch/en.html
University of Copenhagen	https://www.ku.dk/english/
University of Buenos Aires	http://www.uba.ar/internacionales/index.php?lang=en
Tohoku University	http://www.tohoku.ac.jp/en/
National Taiwan University	http://www.ntu.edu.tw/english/
Lund University	https://www.lunduniversity.lu.se/
Durham University	https://www.dur.ac.uk/
The University of North Carolina at Chapel Hill	https://www.unc.edu/
Boston University	https://www.bu.edu/
The University of Auckland	https://www.auckland.ac.nz/en.html
The University of Sheffield	https://www.sheffield.ac.uk/
University of Nottingham	https://www.nottingham.ac.uk/
University of Birmingham	http://www.birmingham.ac.uk/
The Ohio State University	https://www.osu.edu/
Zhejiang University	http://www.zju.edu.cn/english/
Trinity College Dublin, The University of Dublin	https://www.tcd.ie/
Rice University	https://www.rice.edu/
Korea University	https://www.korea.edu/
University of Alberta	https://www.ualberta.ca/
University of Saint Andrews	https://www.st-andrews.ac.uk/
The University of Western Australia	https://www.uwa.edu.au/
The Pennsylvania State University	https://www.psu.edu/
Moscow State University	https://www.msu.ru/en/
The Hong Kong Polytechnic University	http://www.polyu.edu.hk/
University of Science and Technology of China	https://en.ustc.edu.cn/
University of Geneva	https://www.unige.ch/en/university/
Kungliga Tekniska Hogskolan University	https://www.kth.se/en
Washington University in St. Louis	https://wustl.edu/