

Persuasive communication systems: a machine learning approach to predict the effect of linguistic styles and persuasion techniques

Annye Braca and Pierpaolo Dondio

School of Computer Science, Technological University Dublin, Dublin, Ireland

Abstract

Purpose – Prediction is a critical task in targeted online advertising, where predictions better than random guessing can translate to real economic return. This study aims to use machine learning (ML) methods to identify individuals who respond well to certain linguistic styles/persuasion techniques based on Aristotle's means of persuasion, rhetorical devices, cognitive theories and Cialdini's principles, given their psychometric profile.

Design/methodology/approach – A total of 1,022 individuals took part in the survey; participants were asked to fill out the ten item personality measure questionnaire to capture personality traits and the dysfunctional attitude scale (DAS) to measure dysfunctional beliefs and cognitive vulnerabilities. ML classification models using participant profiling information as input were developed to predict the extent to which an individual was influenced by statements that contained different linguistic styles/persuasion techniques. Several ML algorithms were used including support vector machine, LightGBM and Auto-Sklearn to predict the effect of each technique given each individual's profile (personality, belief system and demographic data).

Findings – The findings highlight the importance of incorporating emotion-based variables as model input in predicting the influence of textual statements with embedded persuasion techniques. Across all investigated models, the influence effect could be predicted with an accuracy ranging 53%–70%, indicating the importance of testing multiple ML algorithms in the development of a persuasive communication (PC) system. The classification ability of models was highest when predicting the response to statements using rhetorical devices and flattery persuasion techniques. Contrastingly, techniques such as authority or social proof were less predictable. Adding DAS scale features improved model performance, suggesting they may be important in modelling persuasion.

Research limitations/implications – In this study, the survey was limited to English-speaking countries and largely Western society values. More work is needed to ascertain the efficacy of models for other populations, cultures and languages. Most PC efforts are targeted at groups such as users, clients, shoppers and voters with this study in the communication context of education – further research is required to explore the capability of predictive ML models in other contexts. Finally, long self-reported psychological questionnaires may not be suitable for real-world deployment and could be subject to bias, thus a simpler method needs to be devised to gather user profile data such as using a subset of the most predictive features.

Practical implications – The findings of this study indicate that leveraging richer profiling data in conjunction with ML approaches may assist in the development of enhanced persuasive systems. There are



many applications such as online apps, digital advertising, recommendation systems, chatbots and e-commerce platforms which can benefit from integrating persuasion communication systems that tailor messaging to the individual – potentially translating into higher economic returns.

Originality/value – This study integrates sets of features that have heretofore not been used together in developing ML-based predictive models of PC. DAS scale data, which relate to dysfunctional beliefs and cognitive vulnerabilities, were assessed for their importance in identifying effective persuasion techniques. Additionally, the work compares a range of persuasion techniques that thus far have only been studied separately. This study also demonstrates the application of various ML methods in predicting the influence of linguistic styles/persuasion techniques within textual statements and show that a robust methodology comparing a range of ML algorithms is important in the discovery of a performant model.

Keywords Persuasion, Advertising, Artificial intelligence, Machine learning, Personality traits, Dysfunctional attitudes, Marketing communications

Paper type Research paper

Abbreviations

ANOVA	= Analysis of variance;
AUC	= Area under curve;
AI	= Artificial intelligence;
BalAcc	= Balanced accuracy;
DAS	= Dysfunctional attitude scale;
LDA	= Linear discriminant analysis;
LightGBM	= Light gradient boosting machine;
ML	= Machine learning;
MCC	= Matthews correlation coefficient;
TIPI	= Personality traits;
PC	= Persuasive communication
PCA	= Principal component analysis;
QDA	= Quadratic discriminant analysis;
ROC	= Receiver operating characteristic;
RUS	= Random under sampling;
Sens	= Sensitivity;
Spec	= Specificity; and
SVM	= Support vector machine.

1. Introduction

In the digital information age, crafted messages aimed at influencing how people think are constantly rendered through digital media. This practice is known as Persuasive Communication (PC) and permeates the content of websites, mobile apps, games and social media. Research in the technological domain and in PC are not mutually exclusive, with PC being enriched by its integration with artificial intelligence (AI) and machine learning (ML) approaches. This coalescence presents many opportunities to integrate PC and ML, both to discover more about the effects of PC and to better tailor products and services to individuals. Technologies that currently use PC include AI agents, user profiling and predictive models, with applications in fields such as business, education, health and psychology (Shumanov *et al.*, 2021; Wang *et al.*, 2019; Yang *et al.*, 2019; Zarouali *et al.*, 2022).

PC and the use of linguistic tools of influence, namely, persuasion techniques are well known and extensively studied – for a review, see (Dillard and Pfau, 2002; O’Keefe, 2015; Stiff and Mongeau, 2016). An illustration of one of the most common applications of PC is in sales and marketing. Buyer-seller interactions can be viewed as a PC process, where the

salesperson makes use of persuasive language in an attempt to influence the decisions of a potential buyer. A salesperson using a balanced combination of product knowledge and persuasive language can influence customers to purchase a particular product. As such, the use of persuasive language can be effective in human-to-human sales scenarios (Cialdini, 1987). However, limited research to date has been conducted on the application of persuasive language for human–computer interaction.

Within the e-commerce environment, PC is used in advertising using a range of tools such as Webpage banners, digital nudging (Dennis *et al.*, 2020), website morphing where content is automatically matched to the user’s cognitive thinking (Hauser *et al.*, 2009) or the use of mass defaults as described by Goldstein *et al.* (2008). The latter approach “applies to all customers of a product or service, without taking customers’ individual characteristics or preferences into account”. With some exceptions, e.g. Chen and Lee (2008); D’Souza and Tay (2016); Farseev *et al.* (2021), Shumanov *et al.* (2021), little research has been conducted that uses personality information in PC leveraging ML methods. Current techniques are limited in their ability to communicate with individual customers and could benefit from a more bespoke approach that considers the individual’s personality and character.

The research presented in this paper aims to explore the link between persuasive language style and user engagement, by examining whether variability in linguistic styles/persuasion techniques could affect the level of influence on a given user. To this end, we presented statements with embedded persuasion techniques to more than a thousand survey participants and developed ML classification models to predict the level of influence on a given participant. Model input features included personality traits (Gosling *et al.*, 2003), psychological strengths and vulnerabilities [dysfunctional attitude scale (DAS), (Weissman and Beck, 1978)] and basic demographic information.

The current study adds to a literature that has yielded mixed results with respect to the impact of persuasive messages in the context of human–computer persuasion. To the best of the authors’ knowledge, no prior studies have used the DAS scale as an input for predicting persuasive influence and exploring the individual differences in susceptibility to persuasion. Our hypothesis is that an indicator of psychological strengths and vulnerabilities could help to predict the effect of persuasion techniques given an individual’s profile. Moreover, this study offers a comparison of various persuasion techniques from different lines of research that so far have been studied separately. It also demonstrates the utility of an ML-based approach and the importance of applying a robust methodology that tests multiple ML models in the pursuit of an optimal PC system.

The emphasis of this work is on the computational rather than the psychological side of persuasion; we are primarily interested in investigating if the effect of persuasion can be learnt by a ML model and which features are useful for this task, rather than justifying the observed effect from a psychological and cognitive science standpoint.

The remainder of this paper is organised as follows. Section 2 highlights relevant work. Section 3 provides the methodology used to design the survey, collect and analyse the data. Section 4 presents the results of the ML models. Section 5 provides the discussion. Section 6 concludes the paper.

2. Background

In this section, we provide some theoretical background for the concepts used in our study, and we describe relevant related works. We present the DAS and ten item personality measure (TIPI) scales which were used to collect data relating to personality and emotional

traits of participants. We then introduce the concept of persuasion techniques and describe previous relevant studies where ML techniques or statistical methods were applied to study the relationship between personality and the effect of persuasion techniques.

2.1 *Input features for persuasion systems*

We start by describing the input features used in a persuasion system and the scale we used to collect them. Several investigations have yielded information about the importance of attitude assessment for persuasion. O'Keefe (2015) states that people's attitudes (i.e. values, personality traits, feelings and emotions) are the special interest of a persuader, because attitudes represent stable evaluations that can influence behaviour. Furthermore, the importance of alignment between attitudes and the persuasive message has been emphasised, noting that this functional matching is crucial for the persuasion process (Petty and Cacioppo, 1986). In addition, O'Keefe (2015) argued that attitudes and the belief behind the attitude could determine people's actions, e.g. what products people buy, what policies people endorse, what hobbies they pursue. Furthermore, O'Keefe argues that the systematic study of persuasion requires assessing people's attitudes – with attitudes representing the persuasive target.

With many studies suggesting the importance of assessing and measuring people's attitudes, we reviewed the literature to ascertain useful variables/input features for the development of ML models that can predict persuasive message impact based on user attitude and demographic data. The underlying hypothesis is that the complex relationship between a user's profile and their response to different persuasion techniques can be learnt by ML algorithms. The data used for building persuasive profiling models commonly includes demographic information (e.g. gender, age and postcode), whereas some studies also incorporate personality traits (Shumanov *et al.*, 2021; Matz *et al.*, 2017; D'Souza Tay, 2016b; Anagnostopoulou *et al.*, 2017; Spielmann *et al.*, 2016). A variety of attitude assessment and data collection techniques have been used in studies – for example, data can be supplied voluntarily in exchange for internet services. Furthermore, current trends in research on data collection provide scope for implicit measure assessment. Examples of this type of data collection include digital footprints (e.g. search history, site usage, click data) (Gencoglu *et al.*, 2015), Facebook likes (Marengo *et al.*, 2020; Azucar *et al.*, 2018) and Twitter posts (Setiawan and Wafi, 2020). Data harvested from digital sources is beyond the scope of this paper – for a review, see Laperdrix *et al.* (2020), Pugliese *et al.* (2020). This line of research provides evidence that the quantification of user activity on social media platforms is being explored to measure the effectiveness of targeted persuasive messages (Bossetta, 2018; Farseev *et al.*, 2015; Farseev and Chua, 2017). In this study, participants data were collected via the survey described in Section 3.1. The personality traits and attitude scales are collected using the TIPI and the DAS scale, that we present in the remaining of this section.

2.1.1 Personality traits. The Five-Factor-Model (Gosling *et al.*, 2003) of personality is one of the most heavily used frameworks in academia and industry for gathering personality-based data. The associated characteristics within the Five-Factor-Model can be described as follows:

- (1) *Extroversion.* High levels of extroversion are associated with assertiveness and excitement-seeking. Extroverts tend to be social and outgoing whether in online social communications or face-to-face settings (Yoo and Gretzel, 2011). Conversely, low levels of extroversion are associated with unwillingness to engage in social activities and being inwardly focussed (Hills and Argyle, 2001; Amirkhan *et al.*, 1995).

- (2) *Agreeableness*: High levels of this trait are characterised by pro-social behaviours and a concern for the welfare of others (empathy and altruism). Low levels of agreeableness indicate an excessive or exclusive concern for self-interest and self-advantage (Graziano and Eisenberg, 1997).
- (3) *Conscientiousness*: High levels of conscientiousness suggest high capacity for self-control, regularisation and awareness of one's behaviour and its impact. Low levels of conscientiousness are associated with flexibility, spontaneity and impulsive behaviour (Toegel and Barsoux, 2012).
- (4) *Emotional stability*: High levels of emotional stability (low in neuroticism) are associated with individuals who tend to think before they react emotionally. Features often include resilience, calmness and logical thinking (Vittersø, 2001). Low levels of this trait may result in high levels of anxiety, worry, fear, anger and guilt (Thompson, 2008).
- (5) *Openness to experience*: High levels of open-mindedness are associated with intellectual curiosity and receptivity to new ideas (Ambridge, 2014). Alternatively, low levels in this dimension are linked to apprehensiveness towards novelty and resistance to change even when it may be beneficial. This trait is common in individuals who prefer traditions, routines and familiarity.

In our study, these personality-based features were captured via participants rating themselves based on 10 personality trait descriptions (Likert scale between 1 and 7) in the survey (Appendix). These 10 scores can also be converted into 5 scores by pairing up related traits.

2.1.1.2 Dysfunctional attitude scale. Studies by Quraishi and Oaksford show that the more emotionally based the persuasive strategy, the more the persuasive message could influence individuals' beliefs (Quraishi and Oaksford, 2013). New research by Rocklage and Luttrell indicates that emotions are a predictor of long-lasting attitudes and hence using emotion-evoking persuasive messages can create enduring attitudes (Rocklage and Luttrell, 2021). Contrary to the long-held idea that this kind of influence had a short span, they suggest that this emotional influence can affect the belief behind the attitude, creating a permanent change of opinion. Furthermore, they also argue that "lay individuals generally fail to appreciate the relation between emotionality and attitude stability". Similarly, it has been noted (Bless *et al.*, 1990; Kaptein and Eckles, 2012) that positive emotions (happiness, joy, interest, gratitude, love and contentment) and negative emotions (sadness, anger, loneliness, jealousy, self-criticism, fear or rejection) are considered key factors in the process of influence (Brennan and Binney, 2010; Griskevicius *et al.*, 2009).

Cognitive psychology researchers suggest that a comprehensive measure for emotion assessment is the DAS. Burns (1981) notes that the DAS scale can provide an indicator of the psychological strength and vulnerability of an individual based on their own system of beliefs. DAS positive scores are associated with psychological strength, whereas negative scores represent an area where an individual can be emotionally vulnerable. Additionally, Burns illustrates how the DAS test can reveal negative and positive emotions of an individual and he describes each of the seven areas covered by DAS as follows:

- (1) *Approval*: Positive scores indicate independence and a healthy sense of self-worth, even when confronted with criticism and disapproval. In contrast, negative scores suggest sensitivity to external validation, a tendency to be easily manipulated and vulnerability to anxiety.

- (2) *Love*: Positive scores indicate a healthy sense of love and self-esteem. However, negative scores suggest dependency, desire for attention and the belief that an individual needs to be loved to survive.
- (3) *Achievement*: Positive scores indicate a healthy enjoyment of creativity and productivity without thinking of these as a necessity for life satisfaction or self-esteem. Alternatively, negative scores indicate the potential for a person to be workaholic, their capacity for joy and self-worth dependent on productivity. These individuals may exhibit anxiety related to career failure.
- (4) *Perfectionism*: Positive scores suggest that the person is not compulsively preoccupied, does not fear mistakes, is flexible and open-minded. Comparatively, negative scores relating to this attitude may indicate multiple struggles for a person, due to the pursuit of unrealistic standards. Research suggests that self-critical perfectionism is more likely to lead to negative emotions, such as guilt, distress, anxiety, self-condemnation, procrastination, a tendency to be critical of others, a desire for approval above all else and a tendency to be easily offended by criticism.
- (5) *Entitlement*: Positive scores suggest patience, persistence and an awareness of other people's rights, with no issues accepting others as equals. People with negative scores may possess a sense of entitlement, believing that the world owes them something in exchange for nothing. They may have an expectation of privileges and recognition and a belief that everything that happens should somehow benefit them.
- (6) *Omnipotence*: Positive or negative scores are indicators of the propensity of a person to believe they are the centre of their personal universe and the feeling of being responsible for much of what happens around them. They blame themselves for the negative actions of others who are not really under their control.
- (7) *Autonomy*: Positive scores indicate good self-reliance, while negative scores suggest that a person believes that joy and self-esteem have an external source. This DAS dimension relates to willpower, responsibility and control of one's life.

The DAS assessment of emotions can be applied by asking individuals to score 35 statements (Burns, 1981, p. 271), which helps capture belief-system information, as well as psychological strengths and cognitive vulnerabilities. These 35 questions were included in our study (scored using a Likert scale between 1 and 5). Similar to the personality-based scores, the 35 DAS scores can be transformed into 7 scores with 5 raw statements contributing to each transformed score.

2.2 Persuasion techniques

Persuasion techniques are a set of linguistic features that can modify the intended core statement without necessarily changing the content of information delivered in the message (Holtgraves and Lasky, 1999). Persuasive messages carry embedded persuasion techniques and linguistic strategies such as language style, variability, intensity, tag words, framing and rhetorical devices (Xu and Tan, 2020; Renaldo, 2017; Kaur et al., 2013). These types of persuasive messages aim to elicit the peripheral route of persuasion/System1 (instinctive and emotional thinking) (Kahneman, 2013; Petty and Cacioppo, 1986), rather than the central route (conscious and logical thinking) (Dove, 2021). Persuasive messages are inevitably biased – caused by the rhetorical and linguistic features of persuasive language. When designing persuasive messages, it is common to use relatively informal styles such as imperatives, contractions, clipping and subject/auxiliary omissions (Labrador et al., 2014).

A wide array of persuasion techniques is adopted across a myriad of domains. Researchers often select a technique to use based on the persuasion target or the content of the message, or instead use techniques that are customarily used in their domain. In this paper, we investigate and apply a range of techniques that are standard in fields including psychology, rhetoric and marketing. Persuasive strategies are extensively studied, and much research can be found which includes relevant definitions and descriptions of the persuasion techniques we have used – for a review (Dillard and Pfau, 2002; Myers, 2012; McGuigan, 2011; O’Keefe, 2015; Carey, 1996; Harris, 1997).

In this study, we have used persuasion techniques that elicit the peripheral route of persuasion (also known as System 1) and embedded each technique within a statement relating to the benefit of third-level education. The selected techniques include appeal to finances and logic (Aristotle, 2015; Aristotle and Cooper, 1960), rhetorical devices (anaphora, antanagoge, epistrophe, rhetorical question) (Harris, 1997), cognitive theories (flattery, awareness words, illusion of superiority, priming/semantic priming) (Fogg and Nass, 1997; O’Keefe, 2015; Stengel, 2002) and Cialdini’s Principles of authority and social proof (Cialdini, 1987). A summary with a brief description of persuasion techniques employed in this work is provided in Table 1.

In this research, to acquire the input features for the supervised learning algorithms, a survey was designed (Section 3.1) that captures demographic and attitude information i.e. personality (TIPI) and belief system-based variables (DAS) this is our input matrix. The response variable is a score given by participants which indicates the level to which they felt influenced by a presented statement. The statement contained a core message such as the benefits of education and an embedded persuasion technique. The participant input data and statement scores were then used to develop and evaluate ML models.

Figure 1 shows a proposed design matrix for modelling persuasion. The design matrix consists of two main parts: the independent variables (user profile data) and the dependent variable (the score each user gives to a persuasion technique. Source: Authors work.

2.3 Machine learning

ML is a discipline that models data by integrating and applying methods from fields such as computer science, mathematics, statistics, data mining and distributed systems. The study of the link between users’ profiles and persuasive strategies has been widely confirmed in marketing and advertising research (Hirsh *et al.*, 2012; Clark and Çalli, 2014; Plummer, 2000). However, most of the studies made use of conventional statistical models for predicting the effect of persuasion techniques given the user’s profile (Pangbourne *et al.*, 2020; Thomas *et al.*, 2017; Wang *et al.*, 2019). Also, it has been noted that most studies tend to focus on a few persuasion techniques, for example Cialdini’s principles of persuasion (Cialdini, 1987), disregarding the large array of persuasion techniques available in the literature (McFarland and Dixon, 2019). Although, the interest in using ML in persuasion research is growing exponentially, e.g. (Wang *et al.*, 2019; Lukin *et al.*, 2017a; Shmueli-Scheuer *et al.*, 2019a, 2019b), still there are relatively few proposals that formalise user modelling through ML methods.

Compared to conventional statistical models, ML has the capability of capture nonlinear relationships between input data (for example, personality traits, emotions, values, demographical information) and the associated output (e.g. target linguistic styles/persuasion techniques). A relatively limited number of works in the field of PC and advertising have used ML methods, specifically for the prediction of linguistic styles and persuasion techniques, despite ML methods have been used in multiple fields and industries including experimental psychology and behavioural change science. The adoption of such

Persuasive techniques definitions	
Authority	People tend to attribute greater weight to the opinion of an authority figure, suggesting their views to be more credible. This technique is based on the concepts of credentials, credibility, and history (Cialdini, 1987)
Social proof	This technique is considered a social validation strategy - when individuals observe a group manifesting a belief or behaviour, they are more likely to believe and behave similarly (Cialdini, 1987)
Appeal logic	This rhetorical strategy works by presenting facts that lead people to a specific conclusion. It frames the message using keywords that emphasize features such as facts, evidence, experts, and common sense (Aristotle, 2015)
Finances appeal	Also known as appeal to the hip-pocket nerve this technique makes people feel concerned for their financial wellbeing (Mughan, 1987)
Rht. device anaphora	Anaphora employs repetition of words at the beginning of a phrase. The same word/phrase is repeated initially in two successive sentences (Harris, 1997)
Rht. device rhetorical question	The use of this device stimulates critical thinking and encourages drawing out ideas and underlying presuppositions (Harris, 1997)
Rht. device antanagoge	Antanagoge is used to reduce the impact or significance of what is considered negative. It works by balancing the negative with the positive by placing a positive outlook on a situation that has a negative connotation (Harris, 1997)
Rht. device epistrophe	This is the repetition of the same word or group of words at the end of phrases or sentences. The psychological effect of this device works by giving the impression of certainty in an idea. It encourages recipients to adopt a concept and provokes emotional and psychological attitudes in the audience (Harris, 1997)
Flattery	Flattery as a persuasive technique works by communicating positive things about another person, it appeals to people's vanity without regard of true qualities or abilities (Fogg and Nass, 1997)
Awareness words	Awareness patterns words helps to gain acceptance, bypass resistance, increase responsiveness, embed ideas and suggestions. Typical persuasive cues are: <i>notice, see, realise, aware, experience, discover, consider, contemplate, think about, what if, imagine</i> (Young, 2016)
Illusion of superiority	This is a cognitive bias which arises when one imagines themselves as being superior to the average person along various dimensions, such as intelligence, cognitive ability, and possession of desirable traits (Pietroni and Hughes, 2016)
Priming/semantic priming	Priming technique It helps the audience to see a pattern and be familiar with ideas or words; people typically like things that are familiar to them, the repetition restates and reassure the idea, as a result the audience will pay more attention and remember (Harris, 1997)

Table 1. Persuasion technique descriptions

Sample	Demographic Details	Personality Traits	Belief System (DAS)	Persuasion Technique Target
1	Data _{1,1}	Data _{1,2}	Data _{1,3}	Low
2	Data _{2,1}	Data _{2,2}	Data _{2,3}	Low
3	Data _{3,1}	Data _{3,2}	Data _{3,3}	High
.....
k	Data _{k,1}	Data _{k,2}	Data _{k,3}	High
k+1	Data _{k+1,1}	Data _{k+1,2}	Data _{k+1,3}	High
.....
n	Data _{n,1}	Data _{n,2}	Data _{n,3}	Low

Source: Authors' work

Figure 1. Example of the structure of a training data set with input variables that are based on people's characteristics and attitudes and an output variable which incorporates the response to a persuasion technique

methods more widely in the field of PC would represent a promising line of research with potential to discover more powerful and performant models. A summary of recent publications, methodologies and their goals, along with reported applications is provided in [Table 2](#).

3. Methods

Our aim was to investigate the impact on users of different persuasion strategies embedded in textual statements. The statements were presented to participants via a survey, with the participants designating a score indicating the level to which they agreed with a given statement. Demographic, personality (TIPI) and belief system (DAS) information was collected about each participant and this data was used to develop ML models to predict the effect of persuasion techniques on individuals. Additionally, we followed best practice and the suggestions indicated in ([Anctil, 2008](#); [Damgaard and Nielsen, 2018](#); [Rosenfeld and Kraus, 2016](#); [Strader and Katz, 1990](#)) where authors investigated and designed persuasive statements in an educational context.

3.1 Survey design

To gather the input features for the supervised learning algorithms, a survey was designed that captures demographic and attitude information (i.e. TIPI, DAS scores). To measure the influence of the embedded linguistic styles/persuasion techniques present in statements on participants, a Net Promoter Score ([Reichheld, 2003](#)) was adopted wherein the participant indicates their level of agreement with each statement – survey participants score each statement on a scale from 1 (no effect) to 10 (very convincing). In this study the persuasion context was in the domain of education. Statements contained a core message aimed to persuade about the benefit of pursuing a third level education – with multiple variants being presented that each had an embedded persuasion technique accompanying the core message content. These self-reported scores (called influence scores) formed the response variable representing the ground truth for our supervised ML models. Given that participants may score each statement variant (with different embedded persuasion techniques) differently, any variation in these scores indicate an individual’s relative preference for the linguistic style/persuasion technique.

The survey was deployed using Qualtrics software ([qualtrics.com](#)) in August 2020. It contained four sections, i.e. demographic information, TIPI personality traits ([Gosling *et al.*, 2003](#)), DAS attitude information ([Burns, 1981](#)) and the persuasion tasks (presented statements for rating). The full questionnaire can be found in [Appendix](#). Participants were recruited via the crowd-sourcing marketplace Prolific ([prolific.ac.uk](#)). Inclusion criteria required participants to be native English speakers or non-native with high proficiency. Three attention questions were included in the survey to ensure that participants were engaged. Additionally, participants were provided with instructions requesting them to answer all questions as honestly as possible. They were also given the instruction: “Please answer each of the questions as you feel right now”. The survey was answered by 1,061 participants with data from 1,022 participants ultimately being used. A total of 39 submissions were excluded, as these participants either failed two of the three attention questions, completed the survey in an unreasonably short time or did not fully complete the survey.

To assess the reliability of test scores, we used Cronbach’s α which is an internal consistency estimate that is derived from classical test theory and indicates a theoretical measure of reliability. Internal consistency estimates can be used to measure item homogeneity, i.e. the correlation between items on a test that are intended to measure the

Machine learning approaches toward tailoring persuasion techniques in advertisement and related disciplines, categorized according to respective methodological concept

Context	Aim	Methods	Reported results
Political marketing user models	Prediction of the susceptibility to Cialdini persuasive strategies on liberals and conservative (Demir <i>et al.</i> , 2021)	Data: Survey 195 participants Target: Binary variable (Liberal/Conservative). <u>ML algorithm</u> : Logistic RSVM	Logistic R: 60.4% accuracy SVM: 60.4% accuracy Cubic SVM: 75% accuracy Logistic regression 61%
Political advertising user models	Prediction of the effect of framed political ads from textual data given political attitudes and personality (Zarouali <i>et al.</i> , 2020)	Data: Survey 156 participants Target: Attitudinal response toward political party. <u>ML algorithm</u> : Logistic regression	SVM: 0.73% MAE Active learning: 0.7% MAE
PC campaigns recommender systems	Predicting persuasive strategies targeted at energy saving recommendations, based on users' profile (Sánchez-Corcuera <i>et al.</i> , 2020)	Data: Survey conducted by H2020 European project GreenSoul. Target: Rating scores given to persuasion techniques. <u>ML algorithm</u> : Active Learning, Collaborative filtering, SVM	Hybrid RCNN 74.8% accuracy
Crowdfunding for charities/ recommender systems	Predicting persuasive strategies from persuasive dialogs (Wang <i>et al.</i> , 2019)	Data: Survey-Dialog data set, 1285 participants. Target: Persuasive strategy type. <u>ML algorithm</u> : Hybrid RCNN	Logistic R: channel -text: 50% -Audio 50% -Video 52%
Social influence	Predicting the effect of personality and persuasiveness in social media across different channels (text, audio and video) and the effect of the channel on persuasiveness (Mohammadi <i>et al.</i> , 2013)	Data: Manually annotated videos scripts from movies and YouTube 86 videos (3h -40 mins). Target: Perceived persuasiveness scores. <u>ML algorithms</u> : Logistic regression	BERT model: Reciprocity: 0.81% Scarcity: 0.66% Authority: 0.42% Commitment: 0.44% Liking: 0.52% Consensus: 0.25% CNN: Baseline: 0.629% Tuned: 0.69%
Social engineering: Phishing attacks in social media and advertising	To predict whether a (stand-alone) argument is persuasive or not based on author-reader personality traits from text data (Shmueli-Scheuer <i>et al.</i> , 2019a)	Data: 100 texts from each persuasion technique and a data set resulting in 735 texts. Target Type of persuasion techniques. Algorithm: Pre-training language representations BERT	Bayes classifier: Appeal -Celebrity: 0.48 -Scientific: 0.59 -Society -0.58
Social engineering: user modelling and personalization	To predict whether a (stand-alone) argument is persuasive or not based on author-reader personality traits from text data (Shmueli-Scheuer <i>et al.</i> , 2019a)	Data set: Scraped conversational data from Reddit, 2014-2017: 480K arguments and 56 K authors. Target: persuasive and non-persuasive arguments Algorithm: CNN	
Social engineering: crowd-source opinions	Predicting opinions for specific participants on the believability, convincingness, and appeal of specific arguments (Hunter and Polberg, 2017)	Data set: 50 participants and 30 persuasive arguments. Target: Convincing/Non convincing. <u>Algorithm</u> : Bayes classifier	

Source: Authors' work

Table 2. Publications where PC has utilised ML methods – a brief description of aims is provided, along with specific methods and reported applications

same construct (Henson, 2001). Cronbach's α test was conducted by using the SPSS statistical package (IBM SPSS version 26). For the TIPI data, Cronbach's α was 0.355, suggesting a low level of internal consistency when considering the Vellis scale (Vellis, 2003). Gosling (2017) notes that the TIPI test is expected to possess a low Cronbach's α . Tavakol and Dennick (2011) highlight that Cronbach's α levels are affected by the length of the test – if the test length is too short (as in the TIPI case) the value of alpha is reduced. Comparatively, the DAS test results indicated a high level of internal consistency on our data (Cronbach's $\alpha = 0.786$).

The survey contained sets of statements that were very similar, i.e. the same content, but a different embedded persuasion technique, the presentation of statements was randomised to counteract the impact of repetition of these similar statements on participants. As noted in other studies (Hassan and Barber, 2021; Unkelbach, 2007), repetition could contribute to higher ratings as repeated statements are perceived as more trustful. The persuasive statements (Table 3) have a similar length and were reviewed by two independent researchers and pre-piloted with a sample of 30 respondents before the survey was deployed. The respondents provided feedback on the clarity and intelligibility of the persuasive message. Table 3 shows the list of statements with an embedded persuasion technique, and Figure 2 presents the different phases and processes for our experiments, i.e. the steps involved in the survey development and design support the theoretical foundation of the ML models.

3.1.1 Survey validation: analysis of variance – persuasive statements. The analysis of variance (ANOVA) test is an important tool for researchers studying peoples response's to persuasive ads and banners advertisements (Ku and Chen, 2020; Hussain *et al.*, 2018; Huh and Shin, 2015). A one-way ANOVA was conducted to determine if the means of the influence scores assigned by participants to each statement variant (i.e. a statement with the same core message relating to the benefit of third level education but with a differing embedded persuasion technique) differed significantly. The test was used to check if the participants did react to the persuasive techniques embedded in statements – a significant difference in the means of influence scores across statement variants would indicate that participants reacted differently to persuasion techniques, given that the core content was the same across statements.

The hypotheses to be tested were as follows:

- H0.* There are no differences in the means of influence scores for statements with differing embedded linguistic styles/persuasion techniques; hence, the persuasion techniques did not have any significant effect.
- Ha.* There are differences in the means of the influence scores for statements with differing embedded linguistic styles/persuasion techniques; hence, survey participants were affected by the embedded persuasion techniques.

3.2 Machine learning modelling methodology

Multiple models were developed and tested using the same methodology. Each model uses user profiling data to predict the influence level indicated by a user for a given statement. Statements can carry the same core message but have a different embedded persuasion technique. As such, there is a model for each statement variant. The target variable (ground truth) was initially collected as values from 1 to 10 as users indicated the level of influence that they felt a statement had on them. As used in other studies (Pangbourne *et al.*, 2020; Torgo and Gama, 1996), the effect was modelled as a binary variable indicating a

Statements with embedded persuasion techniques based on: Aristotle means of persuasion, rhetorical devices, cognitive theories of persuasion and Cialdini's principles in the communication context of promoting third level education. Underlined words denote the persuasive cues which form the basis of each persuasion technique within a sentence

Authority	Research shows that a college degree pays off in the long run with data indicating that there is a sizable pay gap between those with a degree and those with a secondary school qualification
Social proof	Most people understand the benefits of earning a good degree. Having a degree may not only help you to land the job you want but you'll have the opportunity to apply for an array
Appeal logic	Given all the advantages of earning a college degree, and learn valuable skills, it is common sense to get a degree - apply for a college degree
Finances appeal	Having a college degree is a necessity. Those who have no college education are at risk of poverty because unemployment rates are higher for those who didn't graduate - poverty hurts, get a college degree
Rht. device anaphora	If you desire a better life, if you desire a better finance and if you desire a better prospect then match that desire with your dedication and discipline. Commit every day and do your best - apply for a college degree
Rht. device rhetorical question	If you study hard, you'll be successful and who does not want to be successful? - apply for college degree
Rht. device antanagoge	Success is not for the weak and uncommitted, champions keep playing until they get it right - apply for a college degree
Rht. device epistrophe	Studying hard will bring success, working diligently will bring success and taking responsibility will bring success - apply for a college degree
Flattery	A person of your intelligence deserves to be successful. Smart people take advantage of their intelligence, get better grades, and go further in college - apply for a college degree
Awareness words	Having a college degree offers security, job satisfaction and higher earning potential. Imagine how confident you will feel having a degree
Illusion of superiority	Only smart people would get this: Education is a key factor for success, take advantage of your intelligence and apply for a college degree
Priming/semantic priming	People who earn a college degree have a better lifestyle, better income, better health care - education offers a good life. Apply for a college degree

Note: The persuasion context is about the benefits of pursuing a college education

Source: Authors' work

Table 3.
List of statements with embedded persuasion techniques used in survey

participant's level of agreement with a statement, i.e. an influence level was specified as low or high. We converted the survey influence scores into a binary variable; 1–5 coded as “0” and 6–10 coded as “1”, indicating a low and high level of agreement, respectively. Once converted, the binary target variable resulted in an unbalanced data set.

Unbalanced data distributions could significantly affect the performance of learning algorithms, equal representation of all the classes in the data set is desired (Zeng *et al.*, 2021), and each model (statement variant) was balanced using the random under sampling technique. This decreases the number of majority class instances by randomly removing data from the original data set (Lemaître *et al.*, 2017). This can help to mitigate bias towards the majority class which may result in poor performance on the minority class and artificially high accuracy estimates (Taghizadeh-Mehrjardi *et al.*, 2020).

A common requirement in ML modelling is data scaling of the input variables. Some ML classifiers will not work without standardisation. In our data set, all input variables including TIPI/DAS variables, reversed-scores and non-reversed-scores were scaled. Additionally, in ML modelling is important to compare the performance of multiple ML

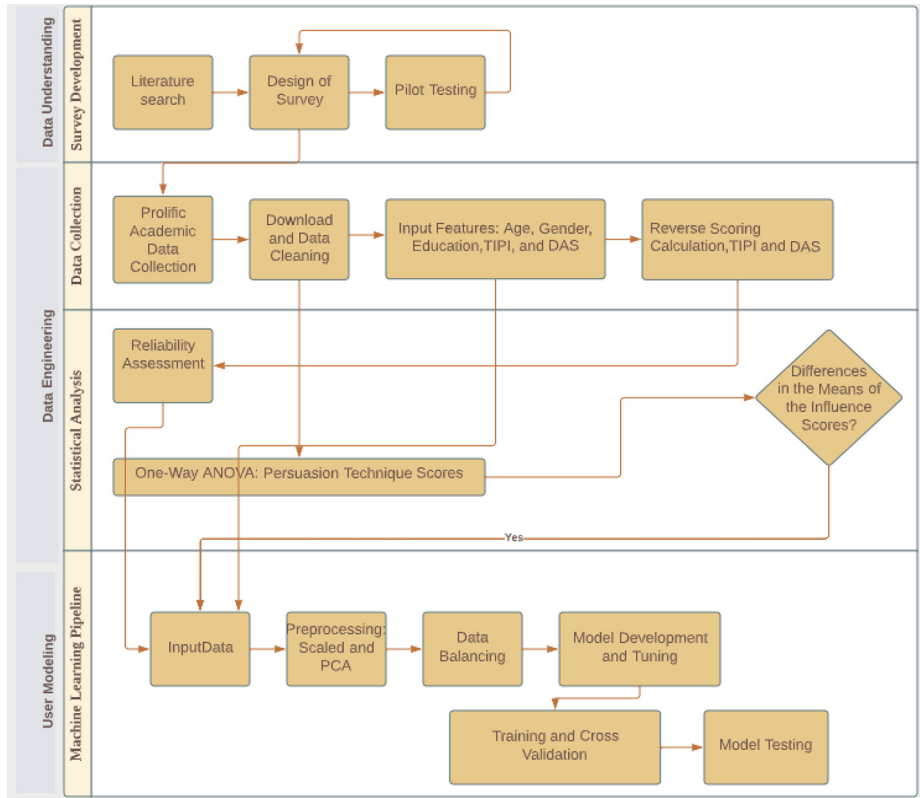


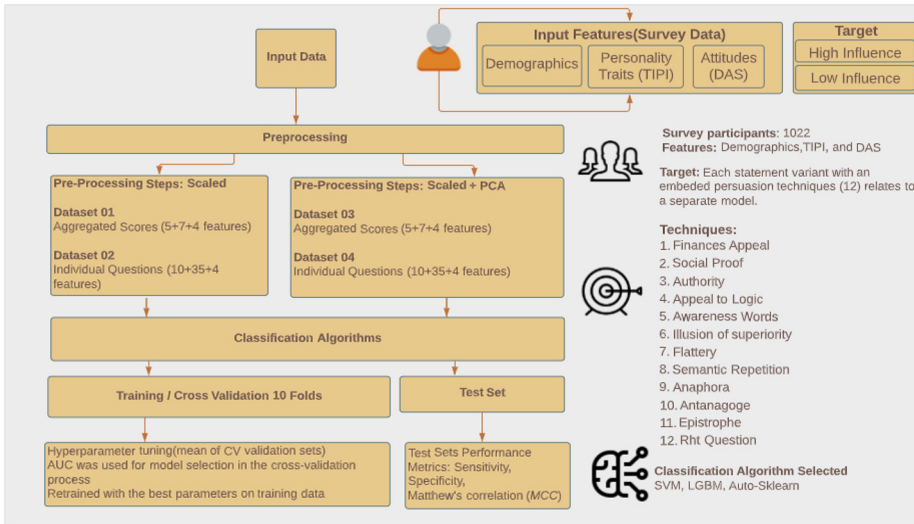
Figure 2.
Data collection involves a process workflow of data acquisition, cleaning and transformation

Notes: The statistical analysis section shows the connection between data processing and important statistical analysis for survey data such as the reliability of the test. The ML pipeline outlines the primary elements of model development and evaluation

Source: Authors' work

algorithms under different pre-processing techniques. PCA is a dimension reduction tool that can dramatically impact the performance of ML models. PCA is used to condense a large set of variables to a small set that still contains most of the information in the large set (Witten and Witten, 2017).

For each of the statement variants investigated, a set of ML models were developed. Three different learning algorithms were tested, i.e. support vector machines (SVMs), gradient boosted machines and auto-sklearn. Using the full data set, training and test sets were formed by random allocation in a 70%:30% split, respectively (Figure 3). Parameter tuning was conducted on the training set to find a good set of model hyper-parameters. This was performed using 10-fold cross-validation with a randomised search of the full parameter search space. Area under the receiver operating characteristic (ROC) curve (AUC) was optimised during the cross-validation tuning process. Final models were then developed using the full training data set and the optimal parameter values discovered in the cross-validation process. These models developed on the full training data were used to make predictions on the independent test set data.



Notes: Variations of the user input data were tested (i.e. reversed-scores and non-reversed-scores). Additionally, the effect of data processing techniques such as scaling and PCA to all versions of TIPI and DAS data. The output variable y (12 unique persuasion techniques) Each model was developed using 10-fold cross-validation on a training split of the total data set. Hyper-parameter tuning was conducted via a cross-validation search on the training set. The retrained model was then used to make predictions on the independent test set

Source: Authors' work

Figure 3. Input data consists of demographic, personality (TIPI) and belief-system information (DAS)

3.2.1 Machine learning features and data set permutation. As alluded to in Sections 2.1.1 and 2.1.2, the survey captures both personality and belief system information (as this type of data is known to affect people's judgement). The personality information is captured via 10 descriptions of personality traits that participants rate themselves on. Examples of these personality traits can be seen in the survey, Appendix, Section 2 with an example being "I see myself as: Sympathetic, warm". The participant then selects a level between "Disagree strongly" and "Agree strongly" on a seven-point Likert scale to best describe themselves for this trait. These scores can be converted to a numeric score between 1 and 7. As such, the 10 personality trait scores form 10 of the input features for the ML models. Similarly, the belief system information is captured via 35 statements (the DAS Scale questionnaire Section 3 of the survey) on which participants rate themselves with an example being "If someone is important to me and expects me to do something, then I should do it". The participant selects a level between "Agree Strongly" and "Disagree Very Much" on a five-point Likert scale to indicate how they think most of the time in relation to the statement. These scores can be converted to numeric values, i.e. 1–5, resulting in 35 numeric input values for belief system-based features.

Aside from age (numeric), the three other demographic features of gender, country of residence and education level were one-hot encoded before inclusion in the ML input matrix. As such, all personality, DAS and demographic data were converted to numeric representations for input to ML models. As already mentioned, the target variable was collected as a score between 1 and 10 as each participant indicated the level of influence that

they felt a specific statement had on them. There were 12 statements (Table 1) presented to survey participants (each with a different core content/embedded persuasion technique pairing); therefore, there were 12 ML modelling tasks, i.e. each with the same input feature data (the same participants were presented with each statement) and different influence scores for each of the 12 statement models.

ML models were developed and tested using the 35 DAS scores, 10 personality scores and the demographic data, as described above. Additionally, however, we wished to investigate models that utilised condensed sets of these features. Specifically, the 35 DAS belief-system scores can be transformed into 7 broad representations – seen in the DAS Scale questionnaire (Section 3) of the survey and described in Section 2.1.2. An example of such is the Omnipotence score which is generated using 5 of the 35 scores i.e. questions 26 through 30 in the DAS Scale questionnaire. As these DAS questions were scored using a five-point Likert scale, the transformation of each set of 5 DAS questions into a single score used a numeric mapping, i.e. depending on the Likert scale response, the following values were mapped: Strongly Agree: -2 ; Agree Slightly: -1 , Neutral: 0 , Disagree Slightly: $+1$, Disagree Very Much: $+2$. The sum of the mapped values for each five-question set (e.g. the five questions relating to Omnipotence) becomes the transformed score. As such, a transformed score for each of approval, love, achievement, perfectionism, entitlement, omnipotence and autonomy is obtained. In the case of the 10 personality-based scores, these can also be condensed into 5 personality scores that represent the traits described in Section 2.1.1. Within the 10 personality statements presented in the survey, there are natural pairs relating to each of the 5 personality traits in the Five-Factor-Model (Section 2.1.1), e.g. questions 1 and 6 relate to Extraversion. The scores given for each pair can be averaged to form a single transformed value which indicates consistency of responses by participants – as a reverse score system is used.

Therefore, in addition to the models developed using the 35 DAS scores, 10 personality scores and demographic data (high dimensional input), models were also tested using the condensed input instead, i.e. the 7 transformed DAS scores, the 5 transformed personality (TIPI) scores and the demographic data (low dimensional input). Additionally, each of the high and low dimensional inputs were processed in two ways, i.e. z -score scaling of features and z -score scaling of features with PCA dimension reduction. As such, with 2 input data set variants (high and low dimensional), 2 different pre-processing approaches and 12 statement models, this constituted 48 model combinations. Given that we tested 3 different ML algorithms, 144 models in total were developed. All models were developed using the same methodology, as described in Section 3.2.

3.2.2 Performance metrics. A range of performance metrics were reported on the test set to compare model efficacies. These included sensitivity (proportion of actual positive cases that were predicted as positive); specificity (proportion of actual negatives correctly identified), F-scores (the harmonic mean of precision and sensitivity), the area under the ROC curve (a measure of how well a classifier performs across all classification thresholds) and balanced accuracy (the average accuracy across all classes) (Brodersen *et al.*, 2010; Kelleher *et al.*, 2020). Additionally, the Matthews correlation coefficient (MCC) was reported which can be a helpful metric when comparing the performance of different ML algorithms, especially when searching for a model that yields similar performance across all classes (Chicco and Jurman, 2020).

Class accuracy is defined by sensitivity accounts for the proportion of people whose influence level were high, and it was classified as high influence; this tells us nothing whether some people whose scores were low would also be classified as high and, if so, in

what proportion. By defining specificity, we addressed the proportion of people whose influence scores were low, and the classifier yielded as low.

3.2.3 Machine learning classification algorithms. Several ML classification algorithms were used for training and testing of our developed models. These included SVMs, gradient boosting machines (LightGBM) and Auto-Sklearn. Briefly:

- *SVM:* This algorithm is used for ranking, classification and prediction in a wide range of applications such as medical applications, weather forecasting and consumer analysis. Training an SVM involves searching for a decision boundary that distinguishes and separates the target feature using a hyperplane [for a review see (Kelleher *et al.*, 2020)]. The tuning parameters include C , γ and the choice of kernel function (e.g. linear, polynomial, radial or sigmoid). The kernel function is used to modify the feature space to make it easier to separate the data set using a hyperplane. C is a penalty parameter related to misclassification while γ relates to the radius of influence of support vectors. To identify the optimal parameters, we performed via a random search process (RandomizedSearchCV in scikit-learn – which function is a speedy algorithm that avoids the combinatorial overload of grid searching by sampling its parameters distribution a fixed number of times as explained by Paper and Paper (Paper and Paper, 2020)). The search parameters used were a logarithmic interval which enabled us to efficiently search a large parameter space. The parameter C was sampled over the interval 10^{-6} to 10^6 , while γ ranged 10^{-8} to 10^8 .
- *Light GBM:* The parameters were tuned during the cross-validation process on the training data set. The first parameter, boosting (GBDT, DART and GOSS) represents the types of gradient boosting methods (Quinto, 2020). The second parameter relates to controlling overfitting by using a leaf-wise tree growth algorithm. The regularisation parameters to tune are the number of leaves. This parameter is one of the most important because it controls the complexity of the model. The third parameter was sub-sample (or bagging fraction), which can improve generalisation and speed of training. It specifies the percentage of rows used per tree building iteration, the rows will be randomly selected for fitting each learner (tree). Forth parameter was feature fraction which refers to column sampling. LightGBM will randomly select a subset of features on each iteration. Fifth parameter, max depth used to limit the maximum depth for tree model. Sixth parameter, bagging fraction which randomly selects part of data without resampling, used only in binary classification and for imbalanced data problem (Ke *et al.*, 2017; Vinayak and Gilad-Bachrach, 2015).
- *Auto-sklearn:* Auto-sklearn (Feurer *et al.*, 2019)) uses 15 classifiers, 14 feature pre-processing methods, 4 data pre-processing methods and a structured hypothesis space with 110 hyperparameters. Additionally, it applies Bayesian Optimization to efficiently navigate the space of possible models and model configurations. The package automatically creates a ML pipeline using a wrapper of the sklearn framework, which includes feature engineering methods and pre-processing techniques (Feurer *et al.*, 2019).

4. Results

4.1 Analysis of variance

An ANOVA analysis was performed (as described in Section 3.1.2) to investigate whether there were significant differences in influence scores for each of the statement variants

furnished to survey participants. Each statement variant relates to a separate ML model. Results of the one-way ANOVA test indicate that the means of the influence scores across statement variants (different embedded persuasion techniques) were significantly different at significance level $p < 0.05$.

The distribution of the means of the influence scores [$F(11,12252) = 86.587, p = 1.028e-189$] suggest that the means of the influence scores were significantly different. Hence, results from ANOVA provide evidence that survey participants perceived the language variation/linguistics style embedded in the statements (Statements – Table 3). Data was normally distributed for each group, as evaluated by Shapiro–Wilk test ($w = 0.999, p = 4.28e-33$). A *post hoc* Tukey test showed that persuasive techniques groups differed significantly at $p < 0.05$.

Results of the ANOVA test support the hypothesis that the means of the influence scores across statement variants (different embedded persuasion techniques) were significantly different, given that the core content was the same across statements. It is important to note that these results do not consider whether survey participants chose to give low or high scores to statements because the context domain of statements or the underlying beliefs. There is evidence to suggest that the content of the message is important in predicting user response. According to Ajzen (2002) and the theory of planned behaviour, persuasion as a process has shown to be dependent on context and on peoples’ beliefs and values. Persuaders can be successful when eliciting specific salient beliefs (attitudes towards the behaviour). As a proof of concept, this study focussed in a very broad topic such as education. For our further studies, we will test an array of persuasion techniques in different communication contexts.

Figure 4 shows the mean distributions of the influence scores for each of the statement with an embedded persuasion technique.

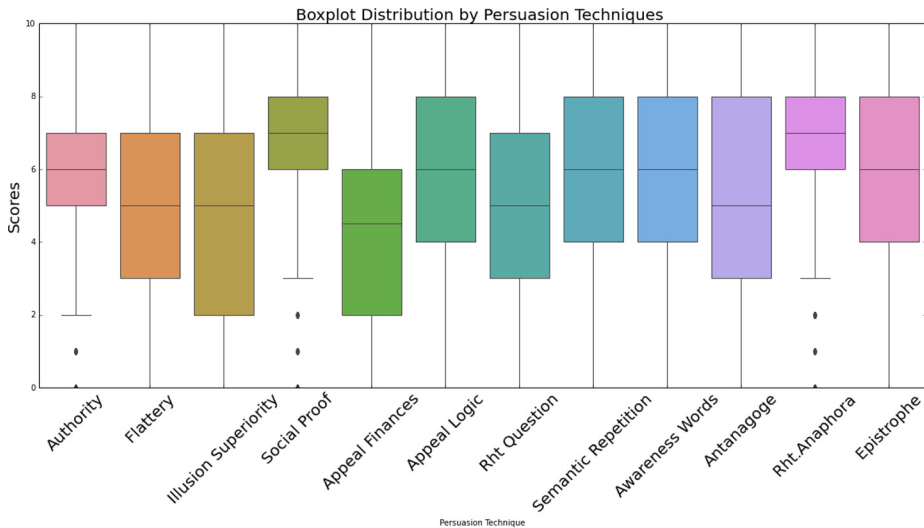


Figure 4. Box plot showing the data distribution of the influence scores by persuasion technique

Note: These included authority, flattery, social proof, priming, anaphora, illusion of superiority, appeal to logic, financial appeal, rhetorical question, awareness words

Source: Authors’ work

4.2 Best models

To assess the quality of predictions on the test set, various metrics were considered to understand each model’s performance. We observed considerable variation in results across the different models i.e. statement models with different embedded persuasion techniques. While some models had an accuracy just above the random guess threshold, others reached reasonable performance (~70% AUC), in line with or better than comparable state-of-the-art works (Wang *et al.*, 2019; Lukin *et al.*, 2017; Shmueli-Scheuer *et al.*, 2019b).

The results in Table 4 show that the best SVM model (antanagoge) achieves generally higher performance than the best LGBM models (Table 5). For example, the best SVM model has an AUC, balanced accuracy, sensitivity and MCC of 0.71, 0.63, 0.72 and 0.27, respectively. Comparatively, the LGBM model (Flattery) which achieves the highest AUC

Technique	AUC	BalAcc	Sens	Spec	MCC	Data set/Pre-processing
Antanagoge	0.71	0.63	0.72	0.55	0.27	Raw data + Scaled + PCA
Anaphora	0.66	0.61	0.78	0.44	0.21	Raw data + Scaled + PCA
Rhetorical question	0.65	0.61	0.71	0.49	0.21	Reversed Scores data Scaled
Flattery	0.65	0.61	0.62	0.59	0.21	Raw data + Scaled + PCA
Appeal finances	0.62	0.61	0.59	0.63	0.21	Raw data + Scaled
Epistrophe	0.61	0.59	0.78	0.41	0.19	Raw data + Scaled
Appeal logic	0.61	0.58	0.49	0.67	0.15	Reversed Scores Scaled data + PCA
Authority	0.58	0.59	0.55	0.62	0.16	Raw data + Scaled + PCA
Awareness words	0.58	0.54	0.63	0.45	0.08	Raw data + Scaled
Illusion of superiority	0.55	0.52	0.55	0.51	0.05	Raw data + Scaled + PCA
Priming-semantic	0.53	0.51	0.72	0.29	0.02	Reversed Scores data + Scaled + PCA
Social proof	0.53	0.54	0.59	0.49	0.06	Raw data + Scaled

Source: Authors’ work

Table 4. SVM results for each statement model – for each technique, the best results of the four permutations of data set and pre-processing are shown

Technique	AUC	BalAcc	Sens	Spec	MCC	Data set/Pre-processing
Flattery	0.66	0.61	0.66	0.54	0.22	Raw data + Scaled
Anaphora	0.65	0.62	0.62	0.59	0.18	Raw data + Scaled
Antanagoge	0.65	0.61	0.64	0.58	0.2	Raw data + Scaled
Epistrophe	0.64	0.61	0.65	0.58	0.22	Raw data + Scaled
Rhetorical question	0.63	0.61	0.65	0.56	0.2	Raw data + Scaled
Authority	0.62	0.59	0.53	0.65	0.18	Raw data + Scaled + PCA
Appeal logic	0.62	0.61	0.56	0.63	0.16	Raw data + Scaled
Finances	0.61	0.58	0.59	0.57	0.15	Raw data + Scaled
Social proof	0.59	0.58	0.59	0.56	0.12	Raw data + Scaled
Awareness words	0.59	0.57	0.58	0.56	0.15	Raw data + Scaled
Semantic repetition	0.56	0.55	0.53	0.56	0.01	Raw data + Scaled + PCA
Illusion of superiority	0.55	0.53	0.53	0.52	0.01	Raw data + Scaled

Source: Authors’ work

Table 5. LGBM results for each statement model – for each technique, the best results of the four permutations of data set and pre-processing are shown

of 0.66 has a balanced accuracy, sensitivity and MCC of 0.61, 0.66 and 0.22, respectively. It is notable that the more automatic approach to ML modelling (Auto-Sklearn) yields the least efficacious results in general. Interestingly, the top 5 models by AUC score for each of the 3 investigated algorithms include the statement models Flattery, Anaphora and Antanagoge. Only 3 of the 36 models outlined in Tables 4–6 use the low dimensional input data representation (i.e. the 35 DAS questions that were condensed to 7 values and the 10 TIPI questions condensed to 5 values). This may indicate that the loss of information when summing and averaging the DAS and TIPI data, respectively, may have limited the models’ ability to find a predictive pattern – given that the majority of models selected used the raw data.

Results pertaining to each ML algorithm (SVM, LGBM and Auto-Sklearn) applied to the various statement models (core message with embedded persuasion technique) are shown in Tables 3, 4 and 5. As explained in Section 3.2.1, there were 48 model combinations for each ML algorithm – here we have shown results relating to the best data set and pre-processing combination (best of four) for each persuasion technique.

4.3 Model performance at high/low influence scores

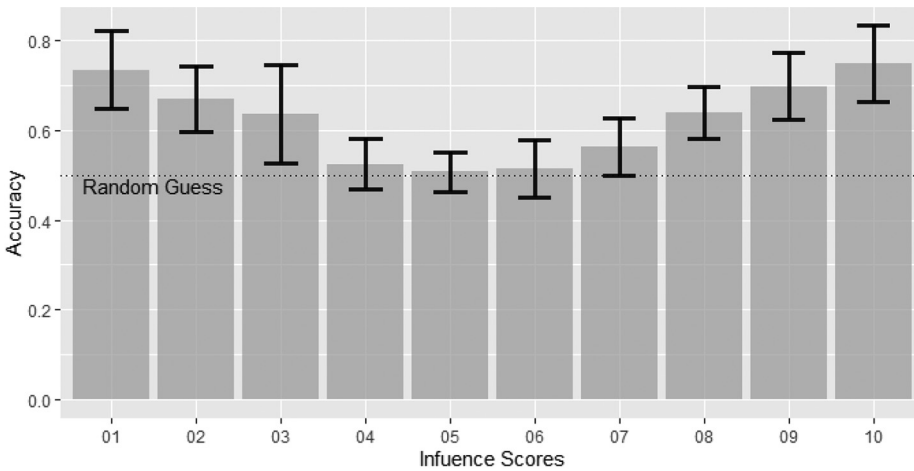
Even though our investigated models were binary classifiers, i.e. they predict a low or high influence score, the underlying discrete influence scores with values ranging 1–10 were available. This allowed us to investigate how sensitive models were at extreme scores (near 1 and 10) and at mid-range scores (around 5 and 6) where the participant’s influence level straddled the low/high binary threshold. As such, we were interested to know to what extent participant’s felt that a statement had a small, medium or strong influence on them.

Figure 5 shows the average accuracy of the 12 LGBM classifiers by influence score (a similar trend is present for SVM). The graph shows how the classifiers obtained good accuracy (approximately 75%) when the effect of persuasion was at the extremes, i.e. little to no effect and high effect (influence scores close to 1 or 10, respectively), while the classifiers performed poorly when the persuasive effect was mid-strength (influence scores around 5 and 6). For the highest-performing techniques such as antanagoge, the classifier reached an accuracy of close to 85% for extreme (high or low) influence scores – for example, this means that in most cases when the actual influence score was 1, the classifier predicted the low influence class and, alternately, when the actual influence was 10, the classifier

Table 6. Auto-Sklearn results for each statement model – for each technique, the best results of the four permutations of data set and pre-processing combinations are shown

Technique	Classifier	AUC	BalAcc	Sens	Spec	Data set/pre-processing
Flattery	Libsvm_svc	0.63	0.55	0.57	0.53	Raw data + Scaled + PCA
Antanagoge	QDA	0.61	0.56	0.62	0.53	Raw data + Scaled + PCA
Anaphora	QDA	0.58	0.54	0.54	0.55	Raw data + Scaled + PCA
Epistrophe	QDA	0.59	0.57	0.54	0.55	Raw data + Scaled + PCA
Appeal finances	Random forest	0.59	0.57	0.59	0.55	Raw data + Scaled
Social proof	Random forest	0.56	0.54	0.61	0.48	Raw data + Scaled
Authority	Random forest	0.59	0.58	0.61	0.55	Raw data + Scaled + PCA
Illusion of superiority	Libsvm_svc	0.56	0.56	0.61	0.53	Raw data + Scaled + PCA
Rhetorical question	Random Forest	0.54	0.54	0.6	0.48	Raw data + Scaled
Appeal logic	Libsvm_svc	0.55	0.55	0.57	0.53	Raw data + Scaled
Awareness words	QDA	0.57	0.51	0.51	0.51	Raw data + Scaled
Priming-semantic	Random Forrest	0.58	0.53	0.63	0.43	Raw data + Scaled

Source: Authors’ work



Notes: When the effect of the persuasion techniques was extreme (scores close to 1 [no effect] or to 10 [highest effect]), the classifiers showed a reasonable accuracy of 75%. When the effect of the persuasion technique was more uncertain (i.e. moderate scores around 5 and 6), the classifiers performance did not differ from a random guess. Error bars of 2 standard deviations are shown

Source: Authors' work

Figure 5. Average accuracy of the 12 LGBM classifiers by influence score

predicted the high influence class. The results shown in Figure 5 are positive, as most of the classification mistakes were minor, occurring in situations where the user indicated a mid-strength reaction to statements/techniques, i.e. there were proportionally less examples of extreme classification mistakes.

4.4 Analysis of features importance

Feature selection is an important process that can achieve a number of objectives including the identification of a subset of variables, the reduction of noise, filtering out irrelevant features while maintaining or even improving the predictive capability of the data (Guyon and Elisseeff, 2003). Feature selection is particularly relevant to the experiment presented in this paper as implementation of an effective persuasion system on a large scale would likely only be feasible using a few highly predictive user data points, e.g. those demographic, personality or belief system questions that have the largest predictive utility. Earlier studies have suggested feature selection can be used to reduce the number of questions that users need to answer, in order to build a useful profile (Kaur et al., 2021).

There are many types of feature selection including filter, wrapper and embedded methods. Additionally, feature transformation methods such as PCA can be used to reduce the dimensionality of data – however these methods do not necessarily yield an explicit set of features and can be less valuable in understanding how features drive a model's decision (Doherty et al., 2022). As such, in this study we have adopted a filter method (LGBM feature importance) to compare the impact of each feature – the importance scores can be used to rank each feature. For each persuasive technique we assigned an ordinal score to the features, giving a score of 6 to the most important feature, 5 to the second and so on, down to a score of 1 for the 6th most important feature for that technique. We then aggregated the scores of each feature across all the techniques.

The ten features (individual questions from the survey) with the highest importance score were the demographics questions (gender and age), the TIPI questions 5 (Openness), 1 (Extroversion) and 4 (Emotional stability) and the DAS questions 18 (Perfectionism), 22 and 24 (Entitlement), 11 (Achievement) and 26 (Omnipotence). These findings confirm that demographic information (gender and age) can have a strong effect on predictions, in line with previous research (Kaptein, 2015). Interestingly, the results show a robust effect of DAS in predictions, providing support for the hypothesis that negative and positive emotions of an individual are key factors of the influence process (as explained in Section 1). Figure 6 shows the importance score by psychological dimension and that many DAS components outperformed the TIPI ones, providing evidence that the information collected by the DAS scale did add value to the predictions. We then grouped the importance scores of each question by the 12 psychological dimensions (5 from TIPI and 7 from DAS). For instance, the Openness score was the average of the score of the two individual TIPI questions 5 and 10. TIPI components are shown in dark grey and DAS components in light grey (see Figure 6).

4.5 Model performance with and without dysfunctional attitude scale information

The AUC averaged over all statement models (i.e. the 12 persuasion technique models) is shown in Figure 7 for both models that used only demographic and TIPI data versus models that used demographic, TIPI and DAS data.

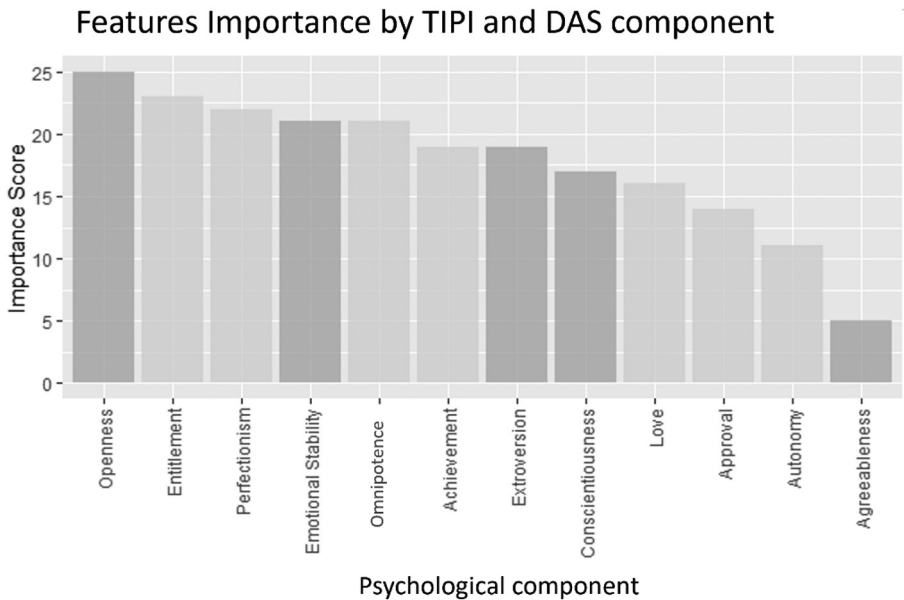
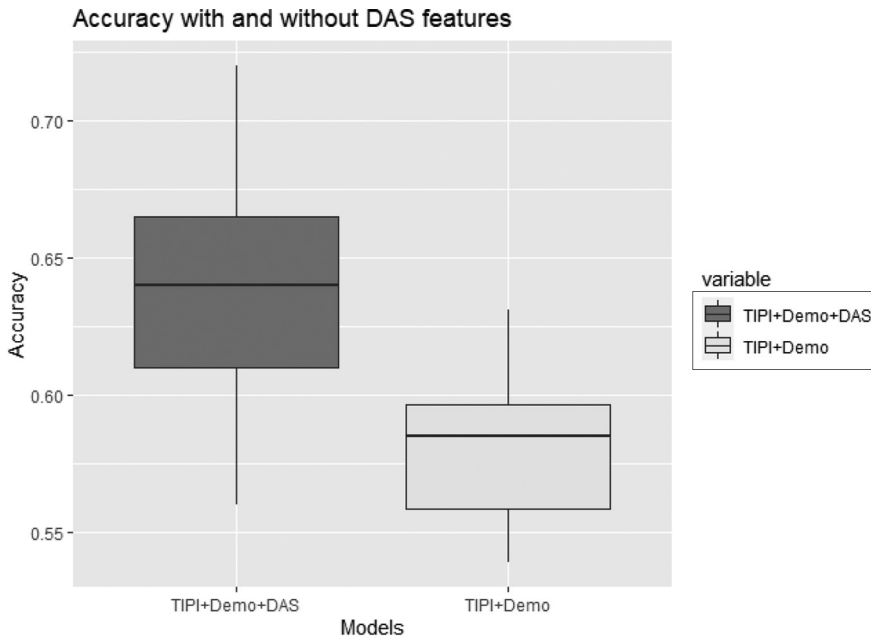


Figure 6.
The importance of each TIPI and DAS psychological dimension

Notes: The importance of each individual question has been aggregated and averaged into the relative psychological dimension (5 for TIPI and 7 for DAS). The TIPI traits are displayed in dark grey and the DAS components in light grey. Entitlement and Openness were the most important features across the 12 persuasion techniques considered in our study

Source: Authors' work



Note: The boxplot shows how the models that include DAS features achieved a significantly better AUC

Source: Authors' work

Figure 7. The average AUC of the 12 models using all features (i.e., TIPI, DAS and demographic, in dark grey) and using only TIPI and demographics (light grey)

5. Discussion

In this study, we investigated the effect of persuasion techniques for different users' profiles, and we evaluated if the effect of the studied techniques can be predicted using ML models. Our results showed that some techniques had an accuracy just above the random guess threshold, while others achieved good performance of approximately 70% balanced accuracy across both binary classes. This is in line or better than comparable state-of-the-art works (Table 2). As a reference point for binary classification performance, models that achieve a balanced accuracy in the region of 50% are no better than guessing. As Kelleher *et al.* noted (Kelleher *et al.*, 2020, p. 399) "a model built to predict which customers would be most likely to respond to an online ad only needs to do a slightly better than random job of selecting those customers that will actually respond in order to make a profit". Therefore, as our best developed models achieved in the region of 70%, this indicates that a signal exists in the user profiling data which can predict high or low influence effects for given statements.

The statements based on rhetorical devices (e.g. antanagoge, anaphora, rhetorical questions) yielded the best performance scores. In the case of SVM learning models, the persuasion statement models that utilised antanagoge generally produced the best performance metric values i.e. AUC of 0.71, balanced accuracy of 0.63, sensitivity of 0.72, specificity of 0.55 and an MCC of 0.27. Similarly, in the case of LGBM learning models, antanagoge was the second-best technique yielding an AUC of 0.65, a balanced accuracy of 0.61 and an MCC of 0.2. Furthermore, the LGBM classification model was one of the best

performing algorithms when applied to the flattery statement model (AUC = 0.66, balanced accuracy = 0.61 and MCC = 0.22). Other statement models that showed good performance included those that incorporated anaphora, appeal to logic and authority embedded persuasion techniques. In the case of Cialdini's principles of persuasion, results indicated moderate performance for the authority and social proof techniques, but these were inferior to rhetorical devices on the investigated data set (Tables 4-6).

Regarding the input data of TIPI personality traits and DAS attitude information, the models trained using these features outperformed those trained without the DAS features. As shown in Section 4.5, the inclusion of TIPI personality traits and DAS dysfunctional attitude information was an important addition to the model - as results showed an increase in classification utility. The top questions from the survey related to DAS and TIPI features (as selected by the LGBM algorithm). Across the 12 statement models, these questions related to perfectionism, openness, entitlement, achievement and emotional stability. This initial attempt at understanding the classifier's decision-making process could be explored further by applying feature selection methods to reduce the number of features (survey questions) necessary for modelling. Selecting a reduced feature set has the potential to improve classification performance and the collection of a small number of features would make any commercial PC system much more viable, e.g. a small set of predictive features (profile data) could be collected from users on registration for online services.

Our results suggest that by leveraging the recent availability and accessibility of ML and deep learning methods, there is an opportunity to advance the field of computational PC. The work in this paper represents a starting point for further experiments - given the favourable results of our ML models, we would further investigate the integration of DAS and TIPI features with an extended range of persuasion techniques. Overall, the observed results are promising and lend weight to the development of user models that can be incorporated into existing advertising platforms for increasing user engagement and potentially sales. One conceivable example involves online advertising systems, either traditional systems or those with customer segmentation capabilities. Often these systems are designed to find a good match between the audience (users) and products or advertisements. The target often is a binary outcome (e.g. clicked or not clicked a button, subscribed, or not subscribed). Kaptein (2015) notes that "persuasive language when improperly targeted can be pointless and at times unfavourable", for example the use of universal marketing text advertisements. Therefore, to have an impact at the individual level, such platforms must be designed to dynamically select the persuasion technique that suits each user in an attempt to influence user response (Kaptein, 2015). Furthermore, McMahan *et al.* (2013) argued that revenue in online advertising is grounded in user response - predictions superior to random guessing have potential to increase the chance of customer engagement and lead to a higher return-on-investment.

Several limitations to our research are suggested. Firstly, there is the need to generalise our ML models to other domains and communication contexts (Zarouali *et al.*, 2022). Relevant work (Ajzen, 1991; Chalaguine *et al.*, 2019; Hadoux and Hunter, 2019) highlights that content of a message is important in predicting user responses and that people choose to believe arguments that align with their preferences. Our work was restricted to one communication context, i.e. education-based messaging, further research and experiments in various communication contexts is required to validate observed models' performance and its generalisation across domains. Secondly, full comprehensibility of ML models and the associated output may be hard to achieve. To mitigate the lack of interpretability of certain ML models, the analysis of the relevance of input features in ML models can provide informative explanations of the underlying decision-making process made by the ML

algorithms. Further research of feature importance would promote transparency and reproducibility.

Furthermore, developing computational PC systems requires expertise in psychology, marketing and ML domains. For example, there are steps such as designing statement variants with embedded persuasion techniques and developing ML models that clearly require an interdisciplinary team effort. Bridging the gap between the ML community and persuasion researchers and enhancing industry–academia collaboration would potentially bring methodological improvements and foster the practical application of these techniques. Another potential limitation relates to the data collection process. In the survey we asked participants to complete a lengthy questionnaire about their psychological and demographic information. Some studies (Meissner *et al.*, 2019; Sherman and Klein, 2021) have argued that self-reported attitudes and values are often in conflict with people’s actual behaviour. Stephens-Davidowitz and Pinker (2017) suggested researchers beware of the social desirability bias - as people wish to look good (even though most surveys are anonymous) survey participants tend to inaccurately report behaviours and thoughts. Bearing this in mind, surveys are the most used research method for data collection in the field of psychology and marketing.

It is important to consider that theories and models of persuasion are not necessarily applicable across all countries and cultures (Morris *et al.*, 2001), and our study was limited to English speaking countries (most respondents resided in the USA, Canada, UK and Ireland). Finally, persuasion and its integration with AI and its methods, promotes robust ethical debate. Many advocates see an opportunity to adopt such technologies to enhance lives and provide social support. However, critics argue that such systems could be a danger to the autonomy of the user, with the potential to sway people’s minds and alter their desires to suit an external agenda. Nevertheless, the integration of PC with AI is expected to increase sharply in the foreseeable future. Given the advances in technology as concerns data collection and generation, in addition to continually evolving ML and AI, the confluence of these developments suggests that research in persuasion communication system design has strong potential for the future.

6. Conclusions

The current study developed ML models to predict the effect of linguistic styles on users given their psychometric profile. Using profiling data as input the models attempt to predict whether statements had a high or low level of influence on participants. Specifically, the study explores the relative performance of linguistic styles/persuasion techniques given collected user information such as demographic data, TIPI personality traits and DAS attitude-related data.

We showed how PC models could be enhanced when the specific message matches the user’s attitudes, values and self-regulatory goals. Our results suggest that users’ attitudes such as openness (TIPI), entitlement (DAS) and perfectionism (DAS) were considered the most important features used by our ML algorithms to predict the effect of persuasion. Additionally, results suggest that traditional rhetorical techniques such as antanagoge, rhetorical question and anaphora had a more predictable effect, outperforming most recent techniques based on social influence. However, rather than concluding that those techniques are better, our conclusion is that techniques based on social influence (e.g. Cialdini’s principles of persuasion) could indeed be more effective when formulated using rhetorical devices. The introduction of input features from the DAS scale significantly improved the performance of the model, showing how such features should be included in a prediction model for persuasion.

Regarding the novelty of this study work, the study is the first to use features from the DAS scale for predicting persuasion effect. Additionally, this study compared various persuasion techniques that insofar have been only studied separately. To the best of the authors' knowledge, there is no ML model equivalent to this work that can be used as a benchmark, and a comparison can be done only with similar models using either a distinct set of features, target variables or research design. We believe this work could encourage additional efforts towards the development and acceptance of PC and ML methods.

Regarding future works, the current ML persuasion models used a data set of 1,022 observations. While this is a good start, part of the future work includes the procurement of larger data set and the definition of a smaller but effective version of the questionnaire used to collect personality traits of users. A higher amount of data will allow to leverage deep learning techniques, that has been shown to uncover complex patterns and learn high-level features in data while outperforming traditional ML algorithms across a variety of applications. It will be of particular importance to further explore different context domains and conduct an exhaustive feature analysis given the persuasion target. The application of ML methods in the field of PC is still in its infancy but moving forward at pace, and we would expect an increase of its relevance in the years to come. We hope this work encourages additional efforts towards the development and acceptance of the integration of PC and ML methods.

References

- Ajzen, I. (1991), "The theory of planned behavior", *Organizational Behavior and Human Decision Processes*, Vol. 50 No. 2, pp. 179-211.
- Ajzen, I. (2002), "Perceived behavioral control, self-efficacy, locus of control, and the theory of planned behavior", *Journal of Applied Social Psychology*, Vol. 32 No. 4, pp. 665-683.
- Ambridge, B. (2014), *Psy-Q: You Know Your IQ-Now Test Your Psychological Intelligence*, Profile Books.
- Amirkhan, J.H., Risinger, R.T. and Swickert, R.J. (1995), "Extraversion: a 'hidden' personality factor in coping?", *Journal of Personality*, Wiley Online Library, Vol. 63 No. 2, pp. 189-212.
- Anagnostopoulou, E., Magoutas, B., Bothos, E., Schrammel, J., Orji, R. and Mentzas, G. (2017), "Exploring the links between persuasion, personality and mobility types in personalized mobility applications", in de Vries, P.W., Oinas-Kukkonen, H., Siemons, L., Beerlage-de Jong, N. and van Gemert-Pijnen, L. (Eds), *Persuasive Technology: Development and Implementation of Personalized Technologies to Change Attitudes and Behaviors*, Vol. 10171, Springer International Publishing, Cham, pp. 107-118.
- Ancil, E.J. (2008), *Selling Higher Education: Marketing and Advertising America's Colleges and Universities*, Jossey-Bass San Francisco, CA.
- Aristotle (2015), *Rhetoric*, in Roberts, W.R. (Ed.), CreateSpace Independent Publishing Platform.
- Aristotle, and Cooper, L. (1960), *The Rhetoric of Aristotle: An Expanded Translation with Supplementary Examples for Students of Composition and Public Speaking*, Prentice-Hall, Englewood Cliffs, N.J.
- Azucar, D., Marengo, D. and Settaggi, M. (2018), "Predicting the big 5 personality traits from digital footprints on social media: a meta-analysis", *Personality and Individual Differences*, Elsevier, Vol. 124, pp. 150-159.
- Bless, H., Bohner, G., Schwarz, N. and Strack, F. (1990), "Mood and persuasion: a cognitive response analysis", *Personality and Social Psychology Bulletin*, Vol. 16 No. 2, pp. 331-345.
- Bossetta, M. (2018), "The digital architectures of social media: comparing political campaigning on Facebook, Twitter, Instagram, and snapchat in the 2016 US election", *Journalism and Mass Communication Quarterly*, SAGE Publications Sage CA: Los Angeles, CA, Vol. 95 No. 2, pp. 471-496.

- Brennan, L. and Binney, W. (2010), "Fear, guilt, and shame appeals in social marketing", *Journal of Business Research*, Elsevier, Vol. 63 No. 2, pp. 140-146.
- Brodersen, K.H., Ong, C.S., Stephan, K.E. and Buhmann, J.M. (2010), "The balanced accuracy and its posterior distribution", *2010 20th International Conference on Pattern Recognition*, IEEE, pp. 3121-3124.
- Burns, D.D. (1981), *Feeling Good*, Signet Book.
- Carey, C. (1996), "Rhetorical means of persuasion", in Ameilie Oksenberg, R. (Ed.), *Esseys on Aristotles Rhetoric*, pp. 399-416.
- Chalaguine, L.A., Hunter, A., Potts, H. and Hamilton, F. (2019), "Impact of argument type and concerns in argumentation with a chatbot", *2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI)*, IEEE, pp. 1557-1562.
- Chatterjee, A. and Basu, S. (2021), "How vulnerable are you? A novel computational psycholinguistic analysis for phishing influence detection", *Proceedings of the 18th International Conference on Natural Language Processing (ICON)*, pp. 499-507.
- Chen, S.-H. and Lee, K.-P. (2008), "The role of personality traits and perceived values in persuasion: an elaboration likelihood model perspective on online shopping", *Social Behavior and Personality: An International Journal*, Scientific Journal Publishers, Vol. 36 No. 10, pp. 1379-1399.
- Chicco, D. and Jurman, G. (2020), "The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation", *BMC Genomics*, Springer, Vol. 21 No. 1, p. 6.
- Cialdini, R.B. (1987), *Influence*, A. Michel Port Harcourt, Vol. 3.
- Clark, L. and Çalli, L. (2014), "Personality types and Facebook advertising: an exploratory study", *Journal of Direct, Data and Digital Marketing Practice*, Springer, Vol. 15 No. 4, pp. 327-336.
- D'Souza, C. and Tay, R. (2016), "Advertising implications and design of messages", *Marketing Intelligence and Planning*, Emerald Group Publishing Ltd., Vol. 34 No. 4, pp. 504-522.
- Damgaard, M.T. and Nielsen, H.S. (2018), "Nudging in education", *Economics of Education Review*, Elsevier, Vol. 64, pp. 313-342.
- Demir, M.Ö., Simonetti, B., Başaran, M.A. and Irmak, S. (2021), "Voter classification based on susceptibility to persuasive strategies: a machine learning approach", *Social Indicators Research*, Springer, pp. 1-16.
- Dennis, A.R., Yuan, L., Feng, X., Webb, E. and Hsieh, C.J. (2020), "Digital nudging: numeric and semantic priming in e-commerce", *Journal of Management Information Systems*, Taylor and Francis, Vol. 37 No. 1, pp. 39-65.
- Dillard, J.P. and Pfau, M. (2002), *The Persuasion Handbook: Developments in Theory and Practice*, Sage.
- Doherty, T., Dempster, E., Hannon, E., Mill, J., Poulton, R., Corcoran, D., Sugden, K., Williams, B., Caspi, A., Moffitt, T.E. and Delany, S.J. (2022), *A Comparison of Feature Selection Methodologies and Learning Algorithms in the Development of a DNA Methylation-Based Telomere Length Estimator*, BioRxiv, Cold Spring Harbor Laboratory.
- Dove, M. (2021), *The Psychology of Fraud, Persuasion and Scam Techniques: Understanding What Makes Us Vulnerable*, Routledge, Abingdon, Oxon; New York, NY.
- Farseev, A. and Chua, T.-S. (2017), "Tweefit: Fusing multiple social media and sensor data for wellness profile learning", *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 31.
- Farseev, A., Nie, L., Akbari, M. and Chua, T.-S. (2015), "Harvesting multiple sources for user profile learning: a big data study", *Proceedings of the 5th ACM on International Conference on Multimedia Retrieval*, pp. 235-242.
- Farseev, A., Yang, Q., Filchenkov, A., Lepikhin, K., Chu-Farseeva, Y.-Y. and Loo, D.-B. (2021), "SoMin. ai: personality-driven content generation platform", *Proceedings of the 14th ACM International Conference on Web Search and Data Mining*, pp. 890-893.

- Feurer, M., Klein, A., Eggenberger, K., Springenberg, J.T., Blum, M. and Hutter, F. (2019), "Auto-sklearn: efficient and robust automated machine learning", *Automated Machine Learning*, Springer, Cham, pp. 113-134.
- Fogg, B.J. and Nass, C. (1997), "Silicon sycophants: the effects of computers that flatter", *International Journal of Human-Computer Studies*, Elsevier, Vol. 46 No. 5, pp. 551-561.
- Gencoglu, O., Similä, H., Honko, H. and Isomursu, M. (2015), "Collecting a citizen's digital footprint for health data mining", *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, IEEE, pp. 7626-7629.
- Goldstein, D.G., Johnson, E.J., Herrmann, A. and Heitmann, M. (2008), "Nudge your customers toward better choices", *Harvard Business Review*, Vol. 86 No. 12, pp. 99-105.
- Gosling, S.D. (2017), "A note on alpha reliability and factor structure in the TIPI | Gosling", available at: <https://gosling.psy.utexas.edu/scales-weve-developed/ten-item-personality-measure-tipi/a-note-on-alpha-reliability-and-factor-structure-in-the-tipi/> (accessed 15 January 2020).
- Gosling, S.D., Rentfrow, P.J. and Swann, W.B. Jr (2003), "A very brief measure of the Big-Five personality domains", *Journal of Research in Personality*, Elsevier, Vol. 37 No. 6, pp. 504-528.
- Graziano, W.G. and Eisenberg, N. (1997), "Agreeableness: a dimension of personality", *Handbook of Personality Psychology*, Elsevier, pp. 795-824.
- Griskevicius, V., Goldstein, N.J., Mortensen, C.R., Sundie, J.M., Cialdini, R.B. and Kenrick, D.T. (2009), "Fear and loving in Las Vegas: evolution, emotion, and persuasion", *Journal of Marketing Research*, SAGE Publications Sage CA: Los Angeles, CA, Vol. 46 No. 3, pp. 384-395.
- Guyon, I. and Elisseeff, A. (2003), "An introduction to variable and feature selection", *Journal of Machine Learning Research*, Vol. 3, pp. 1157-1182.
- Hadoux, E. and Hunter, A. (2019), "Comfort or safety? Gathering and using the concerns of a participant for better persuasion", *Argument and Computation*, IOS Press, Vol. 10 No. 2, pp. 113-147.
- Harris, R.A. (1997), *A Handbook of Rhetorical Devices*, Third Edition, Robert Harris.
- Hassan, A. and Barber, S.J. (2021), "The effects of repetition frequency on the illusory truth effect", *Cognitive Research: Principles and Implications*, SpringerOpen, Vol. 6 No. 1, pp. 1-12.
- Hauser, J.R., Urban, G.L., Liberali, G. and Braun, M. (2009), "Website morphing", *Marketing Science*, Vol. 28 No. 2, pp. 202-223.
- Henson, R.K. (2001), "Understanding internal consistency reliability estimates: a conceptual primer on coefficient alpha", *Measurement and Evaluation in Counseling and Development*, Taylor and Francis, Vol. 34 No. 3, pp. 177-189.
- Hills, P. and Argyle, M. (2001), "Happiness, introversion-extraversion and happy introverts", *Personality and Individual Differences*, Elsevier, Vol. 30 No. 4, pp. 595-608.
- Hirsh, J.B., Kang, S.K. and Bodenhausen, G.V. (2012), "Personalized persuasion: tailoring persuasive appeals to recipients' personality traits", *Psychological Science*, Sage Publications Sage CA, Los Angeles, CA, Vol. 23 No. 6, pp. 578-581.
- Holtgraves, T. and Lasky, B. (1999), "Linguistic power and persuasion", *Journal of Language and Social Psychology*, Sage Publications Sage CA, Thousand Oaks, CA, Vol. 18 No. 2, pp. 196-205.
- Huh, J. and Shin, W. (2015), "Consumer responses to pharmaceutical company-sponsored disease information websites and DTC branded websites", *International Journal of Pharmaceutical and Healthcare Marketing*, Emerald Group Publishing Limited, Vol. 9 No. 4.
- Hunter, A. and Polberg, S. (2017), "Empirical methods for modelling persuadees in dialogical argumentation", *2017 IEEE 29th International Conference on Tools with Artificial Intelligence (ICTAI)*, IEEE, pp. 382-389.
- Hussain, R., Ferdous, A.S. and Mort, G.S. (2018), "Impact of web banner advertising frequency on attitude", *Asia Pacific Journal of Marketing and Logistics*, Emerald Publishing Limited.

-
- Kahneman, D. (2013), *Thinking, Fast and Slow*, 1st pbk. ed., Farrar, Straus and Giroux, New York, NY.
- Kaptein, M. and Eckles, D. (2012), "Heterogeneity in the effects of online persuasion", *Journal of Interactive Marketing*, Vol. 26 No. 3, pp. 176-188.
- Kaptein, M.C. (2015), *Persuasion Profiling: How the Internet Knows What Makes You Tick*, Business Contact Publishers, Amsterdam.
- Kaur, K., Arumugam, N. and Yunus, N.M. (2013), "Beauty product advertisements: a critical discourse analysis", *Asian Social Science*, Canadian Center of Science and Education, Vol. 9 No. 3, p. 61.
- Kaur, H., Poon, P.K.-C., Wang, S.Y. and Woodbridge, D.M. (2021), "Depression level prediction in people with Parkinson's disease during the COVID-19 pandemic", *2021 43rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Presented at the 2021 43rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, IEEE, Mexico, pp. 2248-2251.
- Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., et al. (2017), "Lightgbm: a highly efficient gradient boosting decision tree", *Advances in Neural Information Processing Systems*, pp. 3146-3154.
- Kelleher, J.D., Mac Namee, B. and D'arcy, A. (2020), *Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies*, MIT press.
- Ku, H.-H. and Chen, M.-J. (2020), "Promotional phrases as analogical questions: inferential fluency and persuasion", *European Journal of Marketing*, Emerald Publishing Limited, Vol. 54 No. 4.
- Labrador, B., Ramón, N., Alaiz-Moreton, H. and Sanjurjo-González, H. (2014), "Rhetorical structure and persuasive language in the subgenre of online advertisements", *English for Specific Purposes*, Elsevier, Vol. 34, pp. 38-47.
- Laperdrix, P., Bielova, N., Baudry, B. and Avoine, G. (2020), "Browser fingerprinting: a survey", *ACM Transactions on the Web (TWEB)*, ACM New York, NY, USA, Vol. 14 No. 2, pp. 1-33.
- Lemaître, G., Nogueira, F. and Aridas, C.K. (2017), "Imbalanced-learn: a python toolbox to tackle the curse of imbalanced datasets in machine learning", *The Journal of Machine Learning Research*, JMLR. Org, Vol. 18 No. 1, pp. 559-563.
- Lukin, S.M., Anand, P., Walker, M. and Whittaker, S. (2017), "Argument strength is in the eye of the beholder: Audience effects in persuasion", ArXiv Preprint ArXiv:1708.09085.
- McFarland, R.G. and Dixon, A.L. (2019), "An updated taxonomy of salesperson influence tactics", *Journal of Personal Selling and Sales Management*, Taylor and Francis, Vol. 39 No. 3, pp. 238-253.
- McGuigan, B. (2011), *Rhetorical Devices: A Handbook and Activities for Student Writers*, Prestwick House Inc.
- McMahan, H.B., Holt, G., Sculley, D., Young, M., Ebner, D., Grady, J., Nie, L. et al. (2013), "Ad click prediction: a view from the trenches", *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1222-1230.
- Marengo, D., Montag, C. and Elhai, J.D. (2020), "Digital phenotyping of big five personality via Facebook data mining: a meta-analysis", *Digital Psychology*, Vol. 1 No. 1, pp. 52-64.
- Matz, S.C., Kosinski, M., Nave, G. and Stillwell, D.J. (2017), "Psychological targeting as an effective approach to digital mass persuasion", *Proceedings of the National Academy of Sciences*, Vol. 114 No. 48, pp. 12714-12719.
- Meissner, F., Grigutsch, L.A., Koranyi, N., Müller, F. and Rothermund, K. (2019), "Predicting behavior with implicit measures: disillusioning findings, reasonable explanations, and sophisticated solutions", *Frontiers in Psychology*, Vol. 10, p. 2483.
- Mohammadi, G., Park, S., Sagae, K., Vinciarelli, A. and Morency, L.-P. (2013), "Who is persuasive? The role of perceived personality and communication modality in social multimedia", *Proceedings of the 15th ACM on International Conference on Multimodal Interaction*, pp. 19-26.

- Morris, M.W., Podolny, J.M. and Ariel, S. (2001), "Culture, norms and obligations: cross-national differences in patterns of interpersonal norms and felt obligations toward coworkers", *The Practice of Social Influence in Multiple Cultures*, Lawrence Erlbaum Associates Publishers, Mahwah, NJ, pp. 97-123.
- Mughan, A. (1987), "The 'hip pocket nerve' and electoral volatility in Australia and Great Britain", *Politics*, Taylor and Francis, Vol. 22 No. 2, pp. 66-75.
- Myers, D. (2012), "Social psychology, McGraw-Hill education", available at: <https://books.google.ie/books?id=O1bpXwAACAAJ>
- O'Keefe, D.J. (2015), *Persuasion: Theory and Research*, SAGE Publications.
- Pangbourne, K., Bennett, S. and Baker, A. (2020), "Persuasion profiles to promote pedestrianism: effective targeting of active travel messages", *Travel Behaviour and Society*, Elsevier, Vol. 20, pp. 300-312.
- Paper, D. and Paper, D. (2020), "Scikit-Learn classifier tuning from simple training sets", *Hands-on Scikit-Learn for Machine Learning Applications: Data Science Fundamentals with Python*, Springer, pp. 137-163.
- Petty, R.E. and Cacioppo, J.T. (1986), "The elaboration likelihood model of persuasion", *Communication and Persuasion*, Springer, pp. 1-24.
- Pietroni, D. and Hughes, S.V. (2016), "Nudge to the future: capitalizing on illusory superiority bias to mitigate temporal discounting", *Mind and Society*, Springer, Vol. 15 No. 2, pp. 247-264.
- Plummer, J.T. (2000), "How personality makes a difference", *Journal of Advertising Research, Journal of Advertising Research*, Vol. 40 No. 6, pp. 79-83.
- Pugliese, G., Riess, C., Gassmann, F. and Benenson, Z. (2020), "Long-term observation on browser fingerprinting: users' trackability and perspective", *Proceedings on Privacy Enhancing Technologies*, Vol. 2020 No. 2, pp. 558-577.
- Quinto, B. (2020), *Next-Generation Machine Learning with Spark: Covers XGBoost, LightGBM, Spark NLP, Distributed Deep Learning with Keras, and More*, Springer.
- Quraishi, S. and Oaksford, M. (2013), "6 Emotion as an argumentative strategy", *Emotion and Reasoning*, Psychology Press, p. 95.
- Reichheld, F.F. (2003), "The one number you need to grow", *Harvard Business Review*, Vol. 81 No. 12, pp. 46-55.
- Renaldo, Z. (2017), "Analysis of linguistic features of beauty product advertisements in cosmopolitan magazine: a critical discourse analysis", *TEL-US Journal*, Vol. 3 No. 2, pp. 141-154.
- Rocklage, M.D. and Luttrel, A. (2021), "Attitudes based on feelings: fixed or fleeting?", *Psychological Science*, Vol. 32 No. 3, pp. 364-380.
- Rosenfeld, A. and Kraus, S. (2016), "Providing arguments in discussions on the basis of the prediction of human argumentative behavior", *ACM Transactions on Interactive Intelligent Systems (TüS)*, ACM New York, NY, USA, Vol. 6 No. 4, pp. 1-33.
- Sánchez-Corcuera, R., Casado-Mansilla, D., Borges, C.E. and López-de-Ipiña, D. (2020), "Persuasion-based recommender system ensembling matrix factorisation and active learning models", *Personal and Ubiquitous Computing*, Springer, pp. 1-11.
- Setiawan, H. and Wafi, A.A. (2020), "Classification of personality type based on twitter data using machine learning techniques", *2020 3rd International Conference on Information and Communications Technology (ICOACT)*, IEEE, pp. 94-98.
- Sherman, J.W. and Klein, S.A. (2021), "The four deadly sins of implicit attitude research", *Frontiers in Psychology*, Frontiers, p. 3601.
- Shmueli-Scheuer, M., Herzig, J., Konopnicki, D. and Sandbank, T. (2019a), "Detecting persuasive arguments based on author-reader personality traits and their interaction", *Proceedings of the 27th ACM Conference on User Modeling, Adaptation and Personalization*, pp. 211-215.

-
- Shmueli-Scheuer, M., Herzig, J., Konopnicki, D. and Sandbank, T. (2019b), "Detecting persuasive arguments based on author-reader personality traits and their interaction", *Proceedings of the 27th ACM Conference on User Modeling, Adaptation and Personalization*, pp. 211-215.
- Shumanov, M., Cooper, H. and Ewing, M. (2021), "Using AI predicted personality to enhance advertising effectiveness", *European Journal of Marketing*,
- Spielmann, N., Babin, B.J. and Verghote, C. (2016), "A personality-based measure of the wine consumption experience for millennial consumers", *International Journal of Wine Business Research*, Emerald Group Publishing Ltd., Vol. 28 No. 3, pp. 228-245.
- Stengel, R. (2002), *You're Too Kind: A Brief History of Flattery*, Simon and Schuster.
- Stephens-Davidowitz, S. and Pinker, S. (2017), "Everybody lies: big data, new data, and what the internet can tell us about who we really are".
- Stiff, J.B. and Mongeau, P.A. (2016), *Persuasive Communication*, Guilford Publications.
- Strader, M.K. and Katz, B.M. (1990), "Effects of a persuasive communication on beliefs, attitudes, and career choice", *The Journal of Social Psychology*, Taylor and Francis, Vol. 130 No. 2, pp. 141-150.
- Taghizadeh-Mehrjardi, R., Schmidt, K., Eftekhari, K., Behrens, T., Jamshidi, M., Davatgar, N., Toomanian, N., et al. (2020), "Synthetic resampling strategies and machine learning for digital soil mapping in Iran", *European Journal of Soil Science*, Wiley Online Library, Vol. 71 No. 3, pp. 352-368.
- Tavakol, M. and Dennick, R. (2011), "Making sense of Cronbach's alpha", *International Journal of Medical Education*, Vol. 2, pp. 53-55.
- Thomas, R.J., Masthoff, J. and Oren, N. (2017), "Adapting healthy eating messages to personality", in Oinas-Kukkonen, H., de Vries, P.W., Siemons, L., Beerlage-de Jong, N. and van Gemert-Pijnen, L. (Eds), *Lect. Notes Comput. Sci.*, LNCS, presented at the Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), Springer, Verlag, Vol. 10171, pp. 119-132, doi: [10.1007/978-3-319-55134-0_10](https://doi.org/10.1007/978-3-319-55134-0_10).
- Thompson, E.R. (2008), "Development and validation of an international English big-five mini-markers", *Personality and Individual Differences*, Elsevier, Vol. 45 No. 6, pp. 542-548.
- Toegel, G. and Barsoux, J.-L. (2012), "How to become a better leader", *MIT Sloan Management Review*, Massachusetts Institute of Technology, Cambridge, MA, Vol. 53 No. 3, pp. 51-60.
- Torgo, L. and Gama, J. (1996), "Regression by classification", *Brazilian Symposium on Artificial Intelligence*, Springer, pp. 51-60.
- Unkelbach, C. (2007), "Reversing the truth effect: learning the interpretation of processing fluency in judgments of truth", *Journal of Experimental Psychology: Learning, Memory, and Cognition*, American Psychological Association, Vol. 33 No. 1, p. 219.
- Vellis, R.D. (2003), "Scale development: theory and applications", *Sage London*, Vol. 13, pp. 0.176-1.036.
- Vinayak, R.K. and Gilad-Bachrach, R. (2015), "Dart: dropouts meet multiple additive regression trees", *Artificial Intelligence and Statistics*, PMLR, pp. 489-497.
- Vittersø, J. (2001), "Personality traits and subjective well-being: emotional stability, not extraversion, is probably the important predictor", *Personality and Individual Differences*, Elsevier, Vol. 31 No. 6, pp. 903-914.
- Wang, X., Shi, W., Kim, R., Oh, Y., Yang, S., Zhang, J. and Yu, Z. (2019), "Persuasion for good: towards a personalized persuasive dialogue system for social good", ArXiv Preprint ArXiv:1906.06725.
- Weissman, A.N. and Beck, A.T. (1978), *Development and Validation of the Dysfunctional Attitude Scale: A Preliminary Investigation*, ERIC.
- Witten, I.H. and Witten, I.H. (Eds) (2017), *Data Mining: Practical Machine Learning Tools and Techniques*, 4th ed., Elsevier, Amsterdam.

-
- Xu, H. and Tan, Y. (2020), "Can beauty advertisements empower women? A critical discourse analysis of the SK-II's 'change destiny' campaign", *Theory and Practice in Language Studies*, Vol. 10 No. 2, pp. 176-188.
- Yang, D., Chen, J., Yang, Z., Jurafsky, D. and Hovy, E. (2019), "Let's make your request more persuasive: modeling persuasive strategies via semi-supervised neural nets on crowdfunding platforms", pp. 3620-3630.
- Yoo, K.-H. and Gretzel, U. (2011), "Influence of personality on travel-related consumer-generated media creation", *Computers in Human Behavior*, Elsevier, Vol. 27 No. 2, pp. 609-621.
- Young, S.C. (2016), *Brilliant Persuasion: Everyday Techniques to Boost Your Powers of Persuasion*, Pearson UK.
- Zarouali, B., Dobber, T., De Pauw, G. and de Vreese, C. (2020), "Using a personality-profiling algorithm to investigate political microtargeting: assessing the persuasion effects of personality-tailored ads on social media", *Communication Research*, SAGE Publications Sage CA, Los Angeles, CA, p. 0093650220961965.
- Zarouali, B., Boerman, S.C., Voorveld, H.A.M. and van Noort, G. (2022), "The algorithmic persuasion framework in online communication: conceptualization and a future research agenda", *Internet Research*, Vol. 32 No. 4.
- Zeng, Z., Li, T., Sun, S., Sun, J. and Yin, J. (2021), "A novel semi-supervised self-training method based on resampling for twitter fake account identification", *Data Technologies and Applications*, Emerald Publishing Limited.

Corresponding author

Annye Braca can be contacted at: d18127085@mytudublin.ie

Survey	
Section 1. Questions About Demographics	
1. What is your age?	
2. Which country are you from?	
3. What gender do you identify as?	<ul style="list-style-type: none"> • Male • Female • Other • Prefer not to say
4. What is the highest degree or level of education you have completed?	<ul style="list-style-type: none"> • High School • Bachelor's Degree • Master's Degree • PhD or higher • Trade School • Prefer not to say
Section 2. Questions about personality traits (TIPI)	
Preliminary Information	
For each of the 10 questions below, declare how you see yourself using the following scale:	
1 = Disagree strongly 2 = Disagree moderately 3 = Disagree a little 4 = Neither agree nor disagree 5 = Agree a little 6 = Agree moderately 7 = Agree strongly	
I see myself as:	
1. Extroverted, enthusiastic. 2. Critical, quarrelsome. 3. Dependable, self-disciplined. 4. Anxious, easily upset. 5. Open to new experiences, complex. 6. Reserved, quiet. 7. Sympathetic, warm. 8. Disorganised, careless. 9. Calm, emotionally stable. 10. Conventional, Uncreative	
Section 3. DAS Scale Questionnaire	
Approval, Love, Achievement, Perfectionism, Entitlement, Omnipotence, Autonomy.	
Answering this section is quite simple, put a check in the column that represents your estimate of how you think most of the time. Be sure to choose one answer for each attitude. Because we all are different, there is no "right" or "wrong answer to any statement. To decide whether a given attitude is typical of your own philosophy, recall how you look at things most of the time.	
1. Agree Strongly 2. Agree Slightly 3. Neutral 4. Disagree Slightly 5. Disagree Very Much	
Answer the following questions:	
1. Criticism will obviously upset the person who receives the criticism.	
2. It is best to give up my own interests in order to please other people.	
3. I need other people approval in order to be happy.	
4. If someone is important to me expects me to do something, then I should do it.	
5. My value as a person depends greatly on what others think of me.	
6. I cannot find happiness without being loved by another person.	
7. If others dislike you, you are bound to be less happy.	
8. If people whom I care about reject me, it means there is something wrong with me.	
9. If a person I love does not love me, it means I am unlovable	
10-Being isolated from others is bound to lead to unhappiness	
11. If I am going to be a worthwhile person, I must be truly outstanding in at least one major respect.	
12. I must be a useful, productive, creative person or life has no purpose	
13. People who have good ideas are more worthy than those who do not.	
14. If I do not do as well as other people, it means I am inferior.	
15. If I fail at my work, then I am a failure as a person	
16. If you cannot do something well, there is little point in doing it at all.	
17. It is shameful for a person to display his weaknesses.	
18. A person should try to be the best at everything he undertakes.	
19. I should be upset if I make a mistake	
20. If I don't set the highest standards for myself, I am likely to end up a second-rate person.	
22. It is necessary to become frustrated if you find obstacles to getting what you want.	
23. If I put other people's needs before my own, they should help me when I need something from them.	
24. If I am a good husband/wife (Or partner), then my spouse/partner is bound to love me.	
25. If I do nice things for someone, I can anticipate that they will respect me and treat me just as well as I treat them.	
26. I should assume responsibility for how people feel and behave if they are close to me.	
27. If I criticize the way someone does something and they become angry or depressed, this means I have upset them.	
28. To be a good, worthwhile, moral person, I must try to help everyone who needs it.	
29. If a child is having behavioural difficulties, this shows that the child's parents have failed in some important respect.	
30. I should be able to please everybody.	
31. I cannot expect to control how I feel when something bad happens.	
32. There is no point in trying to change upsetting emotions because they are a valid and inevitable part of daily living.	
33. My moods are primarily created by factors that are largely beyond my control, such as the past, or body chemistry, or hormone cycles, or biorhythms, or chance, or fate.	
34. My happiness is largely dependent on what happens to me.	
35. People who have marks of success (good looks, social status, wealth, fame) are bound to be happier than those who do not.	

Source: Authors work

Figure A1. Survey