Intelligent Systems with Applications 18 (2023) 200201

Contents lists available at ScienceDirect

Intelligent Systems with Applications

journal homepage: www.journals.elsevier.com/intelligent-systems-with-applications

Concepts and experiments on psychoanalysis driven computing

Minas Gadalla [a,*], Sotiris Nikoletseas [a,b], José Roberto de A. Amazonas [c], José D.P. Rolim [d]

[a] Department of Computer Engineering and Informatics, University of Patras, Rion, Achaia, 265 04, Greece
[b] Computer Technology Institute and Press "Diophantus" - CTI, Rion, Achaia, 265 04, Greece
[c] Department of Computer Architecture, Technical University of Catalonia - UPC, Barcelona, 08034, Spain
[d] Department of Computer Science, University of Geneva, Geneva, 1205, Switzerland

ARTICLE INFO

Keywords:
Lacanian discourses
Psychoanalysis computing
GPT-3

ABSTRACT

This research investigates the effective incorporation of the human factor and user perception in text-based interactive media. In such contexts, the reliability of user texts is often compromised by behavioural and emotional dimensions. To this end, several attempts have been made in the state of the art, to introduce psychological approaches in such systems, including computational psycholinguistics, personality traits and cognitive psychology methods.

In contrast, our method is fundamentally different since we employ a psychoanalysis-based approach; in particular, we use the notion of Lacanian discourse types, to capture and deeply understand real (possibly elusive) characteristics, qualities and contents of texts, and evaluate their reliability. As far as we know, this is the first time computational methods are systematically combined with psychoanalysis. We believe such psychoanalytic framework is fundamentally more effective than standard methods, since it addresses deeper, quite primitive elements of human personality, behaviour and expression which usually escape methods functioning at "higher", conscious layers. In fact, this research is a first attempt to form a new paradigm of psychoanalysis-driven interactive technologies, with broader impact and diverse applications.

To exemplify this generic approach, we apply it to the case-study of fake news detection; we first demonstrate certain limitations of the well-known Myers–Briggs Type Indicator (MBTI) personality type method, and then propose and evaluate our new method of analysing user texts and detecting fake news based on the Lacanian discourses psychoanalytic approach.

1. Introduction

User-related and generated data (for simplicity referred to as user data) constitute a core component for social applications based on interactive media technologies. However, in various contexts where user perception is more susceptible to emotional or some other form of bias, user data reliability is often compromised.

A scenario that accurately exemplifies the impact of the lack of reliability of user data can be illustrated in disaster management. In a hypothetical event of a car accident on a frequented highway, where lots of other cars are passing by, we can imagine loads of social media posts describing details about it, before any truly reliable information (e.g. the police arriving) is conveyed. In such a case, if the high volume/uncertain quality user-based information could be instantly filtered, so that a reliable source which e.g. accurately reports the degree of passenger injury severity could be identified, the timely arrival of an ambulance could be of life-saving impact. Of course, the above is a non-trivial task as it implies mechanisms which are yet non-existent; it is certain, though, that an interdisciplinary approach is necessary to capture the diverse aspects of human social way of expressing perception in terms of data and model it in a formal way, in order to conceptualise such mechanisms and infer the reliability of the information and knowledge that can be acquired from such data.

Because of this lack of reliability associated with user data, interactive media technologies-based applications that depend on the characterisation and prediction of the information and knowledge acquired from such data may be severely impaired and are rarely adopted by the actors involved in real life scenarios.

A relevant, recently established research topic is on detecting fake news. In particular, "fake news detection" is defined as the task of

* Corresponding author.
E-mail addresses: gkantalla@ceid.upatras.gr (M. Gadalla), nikole@cti.gr (S. Nikoletseas), jose.roberto.amazonas@upc.edu (J.R. de A. Amazonas), jose.rolim@unige.ch (J.D.P. Rolim).

https://doi.org/10.1016/j.iswa.2023.200201
Received 11 October 2022; Received in revised form 30 January 2023; Accepted 10 February 2023
Available online 15 February 2023

categorising news along a continuum of veracity, with an associated measure of certainty; veracity is compromised by the occurrence of intentional deceptions Conroy et al. (2015). The state-of-the-art methods for detecting the spread of fake news can be coarsely classified into two categories. The linguistic approaches are based on "language leakages" that take place when someone tries to conceal a lie Feng and Hirst (2013). Here, certain verbal aspects are monitored, such as frequencies and patterns of pronouns, conjunctions, and negative emotion word usage; a task that is found very difficult to achieve. On the other hand, the network approaches are based on corresponding properties and behaviour of how the news is spread. Here, linked data and social network behaviour are studied Ciampaglia et al. (2015).

We conjecture that psychoanalysis theories may be used to provide the tools for a third methodology to be developed. One may choose to disseminate fake news for several reasons; e.g., due to being irrational or because there is something to gain. Independently of the different motives, certain text qualities characteristic of fake news can be captured by a psychoanalytic examination of the texts.

The overarching aim of this research work is to develop a radically new theoretical framework for interpreting user-generated data in the context of social interactions. In particular, by combining elements from two very divergent disciplines – Computer Science and Psychoanalysis – this work will develop the theoretical methods and tools for gaining a deep, holistic understanding of the behavioural context of individuals, groups and crowds from the data they generate. This work focus on improving and providing a fundamentally new perspective **in terms of the corresponding technologies**.

In this context, this research has the ambitious goal of laying the foundations for a new paradigm of **psychoanalysis-driven technologies**.

The specific contributions of this paper, within the broader aims stated previously, may be summarised as follows:

- evaluation of the fake news detection accuracy method based on the personality traits concept demonstrating its limitations;
- proposal of a new method to classify user data based on the Lacanian discourses psychoanalytical concept;
- evaluation of the fake news detection accuracy based on the Lacanian discourses approach;
- definition of a framework and roadmap for the future development of psychoanalysis-driven computing.

After this Introduction, Section 2 describes and compares related works published in the recent years, Section 3 describes the personality traits concept and introduces the novel psychoanalysis-driven approach based on Lacanian discourses, Section 4 evaluates the potential of the adopted Psychological and Psychoanalytic approaches to identify reliability related characteristics of enunciations, Section 5 presents the computational approach followed, Section 6 discusses the obtained results and Section 7 summarises the conclusions and proposes a roadmap for future work.

2. Related work

Relevant, yet different, approaches of combining psychological and social dimensions with computational methods include computational psycholinguistics, personality traits, behavioural analysis, emotional states and cognitive psychology methods. Compared to all those approaches, our approach is fundamentally different since we adopt a psychoanalytic perspective; in particular, we employ the powerful notion of Lacanian discourse types. To the best of our knowledge, this is the first attempt of systematically bringing together psychoanalysis and computing. We believe that such a psychoanalytic approach is eventually more effective compared to the previously mentioned methods, since it addresses deeper, fundamental elements of human personality, behaviour and expression which usually escape methods operating at a "higher" conscious layer.

Having stressed this general novelty of our research methodology, we below discuss relevant, recent research related to the particular case study (fake news detection) which we use in order to exemplify our method.

A fundamental approach of combining psychology with computational linguistics (based on abstract formulations of phrases via a collection of finite-state transition networks) is described in Kaplan (2019); in particular, the author envisions sophisticated natural language technologies as a key factor for improving the (rather poor) performance of current conversational systems used by modern technology. The abundant availability of massive data along with effective AI methods (including deep learning) is expected to further facilitate this vision. We note that, although the notion of conversation is directly relevant to the notion of discourse, the approach taken in that paper is more limited (psychological) than the directly psychoanalytic attempt we pursue in our research.

For the more concrete aspect of detecting misinformation in online social networks, Kumar and Geethakumari (2014) suggests the application of cognitive psychology concepts. An efficient algorithm for detection of spread of misinformation in Twitter is proposed, based on text and network-wide qualities such as the consistency of message, the coherency of message, the credibility of source and the general acceptability of message in the network. In Sastrawan et al. (2022) the authors using deep learning methods achieved over 97% accuracy on detecting fake news in some fake/real news datasets. Again, no psychoanalytic elements are considered when evaluating the qualitative properties of the text. Also, the use of objective, global information is employed, in contrast to our approach which focuses on each text separately (however, our method can also be extended to include global information about texts).

Psychological factors (in particular, emotions as expressed in Reddit conversations) are addressed in Guo et al. (2021), to propose a model for passively detecting mental disorders. The suggested model is based entirely on emotional states and the transitions between these states identified in Reddit posts, in contrast to content-based representations (e.g., n-grams, language model embeddings etc.) in the relevant state of the art. The scope is to overcome the domain and topic bias of content-based representations, towards more general applicability. In a different application domain Chen et al. (2021), the one of intelligent selling, customer personality profiling has been employed; in particular, a psychology based approach utilizing personality traits is taken, without however a psychoanalytic dimension. Our approach aims to avoid a content-specific bias, focusing on underlying qualities of texts captured by the discourse type identification. In fact, discourse types are an even more "primitive" aspect of texts, more so than emotions; so the generality of our method might be broader.

In another psychology-based research for fake news detection, a behavioural analysis approach is taken Cardaioli et al. (2020). In particular, the authors use supervised learning algorithms to profile fake-news spreaders, based on the combination of Big Five personality traits and stylometric features. The method is evaluated on a tweeter dataset in both English and Spanish languages. In a similar spirit, Sampat and Raj (2016) aims to understand the motivations for sharing fake news and the corresponding personality traits among social media users in India; in particular, the findings suggest that the passing time, information sharing and socialisation gratifications lead to instant sharing news. Also, people who exhibit extraversion, neuroticism and openness share news on social media platforms instantly; in contrast, agreeableness and conscientiousness personality traits lead to authenticating news before sharing. We note that that work focuses on fake news spreading, not fake news identification itself. Also, as stated previously, although those works exhibit a psychology-based flavour too, our own research is based on psychoanalytic methods, particularly for identifying the power of characteristic Lacanian discourse types to identify specific charac-

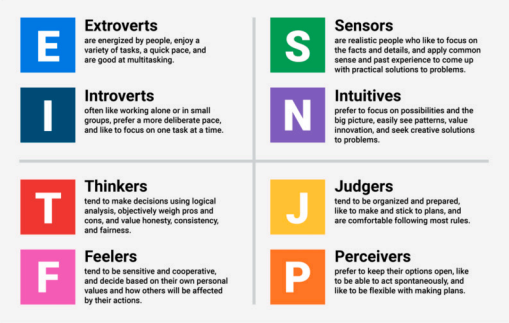

Fig. 1. Personality types groups.

teristics in texts in interactive media as, for example, the fake news identification that we selected to develop the case study of this work.

3. Psychological and psychoanalytic approaches taken

In this section, we describe two distinct approaches to potentially infer reliability-related information from user data. The first one is based on the psychological concept of personality traits, and the second one on the psychoanalytical concept of Lacanian discourses.

3.1. Personality type prediction

Two potential answers to the question "What makes people unique and different from one another?" are motives or traits John and Robins (1994). Regarding motives, we can find the work "Motivation and personality: Handbook of thematic content analysis" Smith (1992) and Murray's list of needs Murray and McAdams (2007). However, no computational approach has been implemented or proposed. Personality traits and types, on the other hand, have numerous theories, taxonomies, and frameworks available, of which the most well known are the Big Five Rothmann and Coetzer (2003) and the Myers–Briggs-Type-Indicator (MBTI) Boyle (1995). More specifically, the trait approach, rather than the motives approach, is still the most widely used and most accepted conceptual framework for describing human personality and successfully predicting human behavior, and has been implemented and applied multiple times in computer-based applications, such as in social media advertising Clark and Çallı (2014), trait prediction from facial images Kachur et al. (2020), and many others. Both these systems are intended to indicate or predict the subject's personality, and thus his or her behavior and preferences, by breaking them down into many dimensions, or groups. In this research, we are focusing mainly on the traits approach and particularly the MBTI system.

The Big Five personality traits comprise the following: extraversion, conscientiousness, agreeableness, openness to new experiences, and neuroticism Rothmann and Coetzer (2003). Meanwhile, the MBTI personality model describes an individual's preferences/behaviour in four dimensions/groups which are also illustrated in Fig. 1:

- Extraversion-Introversion: This dimension measures how a person gets their energy. Extraverts are energized by being around people, while introverts are energized by being alone.
- Sensing - Intuition: This dimension measures how a person takes in information. Sensors rely on their five senses to gather information, while Intuitives rely on their gut feelings and intuition.
- Thinking- Feeling: This dimension measures how a person makes decisions. Thinkers use logic and reason to make decisions, while Feelers use their emotions and values.
- Judging - Perceiving: This dimension measures how a person lives their life. Judgers like structure and order, while Perceivers are more flexible and spontaneous.

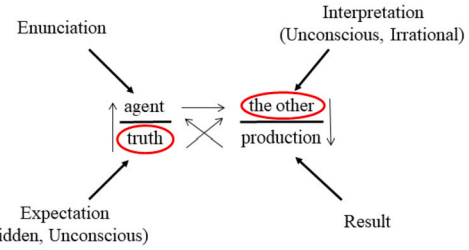

Fig. 2. General representation of the Lacanian discourses.

This way, each individual is classified in terms of one of 16 possible four-letter codes, like ESFJ, indicating a person who may:

- Extraverted (E): is more concerned with the outside world of people and things than with the inner world of ideas.
- Sensing (S): would rather work with known facts and solid experience than explore possibilities or meanings.
- Feeling (F): steps into situations to weigh human values and motives. Prefers to make decisions on the basis of values.
- Judging (J): prefers a planned, determined orderly way of life to a flexible spontaneous.

MBTI was developed by Isabel Briggs Myers in the 1940s Boyle (1995) to implement Jung's Jung (1974) theory of psychological types and the results from such an approach are obtained classically from test booklets, answer sheets, and score keys, and are produced professionally Cattell (1943) Boyle (1995). In Sections 4 and 5 we discuss and use some computational ways of determining personality types using semantic and linguistic analysis and present the results from a Fake/Real news detection application.

3.2. Lacanian discourses

The Four Discourses theory constitutes an attempt of formalisation of the different ways people relate to each other and the economy of knowledge and enjoyment in social relationships. The Lacanian framework defines a more complex representation of the roles assumed by two interacting parties, formulating four discrete discourse types Lacan (1972), Lacan (1974) and Bailly (2009): **Discourse of the Master** – Struggle for mastery/domination/penetration; **Discourse of the University** – Provision and worship of "objective" knowledge – usually in the unacknowledged service of some external master discourse; **Discourse of the Hysteric** – Symptoms embodying and revealing resistance to the prevailing master discourse; **Discourse of the Analyst** – Deliberate subversion of the prevailing master discourse.

Later, Lacan defined an additional fifth discourse, which is not considered in this work, the discourse of the Capitalist Contri (1978), where the subject is commanded to enjoy in the form of commodities.

Fig. 2 presents the general representation of the Lacanian discourses. The four places of the discourse: the "agent", the giver of the discourse; the "other", the one to whom the discourse is addressed; under the message of the agent is hidden the "truth", which is masked by the official statement; and hidden under the other is the "production", or what the agent gets out of the relationship. It is important to realise that the influence of the hidden "truth" upon the "agent" and the enunciation interpretation performed by the "other" are subjective unconscious processes of the communicating parties.

It is possible to draw a parallel between the terms of a discourse and the components of a communication process, in such a way that the dynamics of a given discourse, i.e., the internal relations between elements arranged in different places, can serve to characterise the dynamics of a given media process.

The elements of a discourse or the roles assumed by the communicating parties are the following:

- **S1**: the master signifier represents the true essence of the subject. It may be summarised by *who I am*.
- **S2**: represents the knowledge of the subject. It may be summarised by *what I know*.
- **a**: represents the object cause of desire. It may be summarised by *what I want*.
- **$**: represents the barred subject castrated by the language. It may be summarised by *what I speak*.

The discourses are defined by the position each element occupies in the general representation depicted in Fig. 2. Each discourse representation is called, in the Lacanian jargon, a *matheme*. The definitions given below are reproduced from Bailly (2009).

$\left\uparrow \frac{S1}{\$} \quad \overline{\times} \quad \frac{S2}{a} \right\downarrow$ **Master Discourse:** the master signifier (**S1**) (of the Master) is the agent of communication and instead of addressing the other (**a**), addresses the "knowledge" (**S2**) in his/her place; in other words, the Master is addressing the other not as a Subject but in his/her functional role, because of his/her ability or knowledge (S2). The truth of the master signifier is actually a barred Subject (**$**) and, as lacking as everyone else, is hidden; however, beneath that master signifier (**S1**), the barred Subject (**$**) is enjoying the production of the knowledge (what comes from **a**).

$\left\uparrow \frac{S2}{S1} \quad \overline{\times} \quad \frac{a}{\$} \right\downarrow$ **University Discourse:** the discourse of the University makes a point about the functioning of institutions, and by extension of the individuals within them in their capacity of incarnating the institution. Knowledge (**S2**) occupies the place of the agent, which addresses itself to the object cause of desire (**a**), as the desire for knowledge is the supposed reason why the student is there. However, in this relationship, one can see that the *objet petit* **a** is also, and perhaps as importantly, fed into by the master signifiers of the institution (**S1**), and these contribute endlessly to the castratedness of the Subject of the student. In addressing the knowledge not to the Subject, but to the object cause of desire of the Subject, what is "produced" in, in fact, more castratedness (**$**). Beneath the appearance of dispensing knowledge (**S2**), the University controls the Subject (**$**) by means of its master signifiers (**S1**) and enjoys the "production" of the castrated student (**$**). The institution is also guilty of giving the impression to the student that by careful attention and absorption of its master signifiers, she/he may overcome his/her castration. This is a system of functioning that is common to all institutions: corporations, professions, and government departments, indeed in any institution where "knowledge" (**S2**) in some form takes the place of the agent which addresses discourse, and acts as a lure to the others desire (**a**).

$\left\uparrow \frac{a}{S2} \quad \overline{\times} \quad \frac{\$}{S1} \right\downarrow$ **Analyst Discourse:** in the discourse of the Analyst, the Analyst has accepted becoming symbolically, the *objet petit* **a** of the analysand. This is one of the most usual roles the analyst has to accept; he/she is an empty mirror upon which everything may be reflected, and when in full transference, the analysand (**S1**) will be addressing his/her object cause of desire (**a**). In this case, in his/her role as the *objet petit* **a**, the Analyst addresses his/her discourse to the castratedness (**$**), the anxiety, of the patient, and his/her questioning pushes the analysand to produce a master signifier (**S1**) which is reflected back to the Analyst, while the hidden knowledge of the Analyst (**S2**), in the place of truth, is fed into the castrated Subject (**$**), fuelling the production by it of master signifiers (**S1**). The analysand will discover that the knowledge of its own desire (**S2** hidden beneath **a**) is not held by the Analyst but revealed through its master signifier (**S1**). The Analyst does not adopt a position of power like the Master, or of knowledge, like the University, and because of that, is often considered subversive by institutions.

$\left\uparrow \frac{\$}{a} \quad \overline{\times} \quad \frac{S1}{S2} \right\downarrow$ **Hysteric Discourse:** one doesn't have to be hysterical in the clinical sense to hold the Discourse of the Hysteric; indeed, Lacan made it clear that this type of discourse in non-hysterical people, is precisely what leads to true learning. The agent of the discourse is the castrated (**$**) shortage of the Hysteric; hidden beneath its bar is his/her object cause of desire (**a**). This barred Subject (**$**), driven by its *objet petit* **a**, addresses the master signifiers of the other (**S1**), which respond with the production of knowledge (**S2**), beneath the bar. It is to the master signifier (**S1**) that the Hysteric (**$**) addresses his/her questions, but she/he receives as an answer only the knowledge (**S2**) of that person, which the Hysteric enjoys for want of anything better, although these answers never constitute a satisfactory response to his/her desire (**a**). The Discourse of the Hysteric is held by anyone who is on the path to knowledge. It is a position that requires perfect acceptance of one's ignorance, no great desire to pretend to any other status, and a hunger for the object cause of desire.[1]

Precisely, the idea is that **by distinguishing the discourse type, the ≪ truth ≫ status of an enunciation can be qualified**; i.e., it is the formal characteristics of the discourse which will inform if the truth is based on authority, on documented sources, on the needs to identify the essence of the other, or on the needs to dig for information, respectively. **The formal characteristics will be a key to measure the "reliability" likelihood of the enunciation**. While the personality traits tell us about the speaker, the Lacanian discourses tell us about what is said. A trained psychoanalyst can detect the above signifiers based on semantical analysis of the language.

It can be argued how the Lacanian discourse approach can be extended to groups with more than two interacting parties. The answer to this concern comes from recognising that psychoanalysis is a general framework for the interpretation of situations expressed in any format and by any number of people. Since its development by Freud it has been stated and shown that it can be used to interpret works of art Freud (1914), to analyse social situations Freud (1921), Freud (1933) or to conjecture about the future of civilisation Freud (1927), Freud (1930). These are just a few examples of Freud's works applying Psychoanalysis techniques out of the psychoanalytical setting of a patient and an analyst.

Lacanian discourses are a formal framework to apply psychoanalytical concepts to interpret any real life situation. As the aforementioned Freud's works, they can be used as a reference model to interpret works of art, social situations, and so on.

Nevertheless, the association of an interaction with one of the discourses is not an easy task and has to be dealt with caution. An example that illustrates how a single media phenomenon can be seen from various angles according to the discourse theory is provided by Castro (2016): "When Google scans the Internet collecting information from each site, we are in the discourse of University. When it meets our demand for providing results, we are in the discourse of hysteria. When we deify it, we are in the discourse of the master. When it computes our data and customises the results it offers us, as if it knew us, knew our preferences and anticipate what we want, we are in the discourse of capitalism". Therefore, it is paramount that the context is well defined before proceeding to association, so that the stakeholders of the discourse are clearly identified.

Context depends on the kind of application. The knowledge of the application implies the knowledge of the context and establishes the contextual elements to be considered.

[1] It is a characteristic of the Hysteric discourse to be charged with emotion.

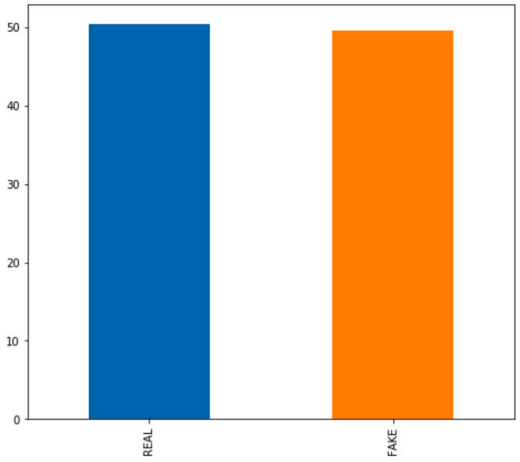

Fig. 3. Label Real = 0 and Fake = 1 distribution.

Table 1
Three random selected headlines from the dataset for illustration purposes.

Headline	Label
Police Turn In Badges Rather Than Incite Violence Against Standing Rock Protestors	1
Republican race enters a new volatile phase	0
First Take Wall Street bids goodbye to June hike	0

4. Preliminary evaluation of the psychological and psychoanalytic approaches

In this section, we empirically evaluate the potential of the Psychological and Psychoanalytic approaches described in Sections 3.1 and 3.2 to identify reliability related characteristics of enunciations.

4.1. Dataset used for evaluation

The selected dataset was obtained from Kaggle[2] and consists of a number of news headlines and content that have been used to develop Real-Fake news algorithms as, for example by Koury and Hernandez (2019); they obtained a 97% accuracy of Fake/Real news prediction utilizing the headlines and content, using the state-of-the-art NLP methods. In our approach, from this dataset, only the headlines were used and submitted to the personality traits assignment algorithm developed by.[3]

Each headline also has an associated label. When the label is equal to 0, the corresponding news is Real, otherwise it is Fake.

For the needs of this application after analyzing the content we decided to remove any headline which was less than 4 words, as most of them don't give any semantic information if it is fake news or not. The final dataset, containing 5860 unique *headlines*, is balanced, ideal for the task of applying the MBTI theory and the labels are equally distributed as shown in Fig. 3. Finally, for illustration purposes, Table 1 contains three random selected rows from the dataset.

4.2. Evaluation of the personality traits (MTBI) approach

An algorithm was developed in the R language to evaluate the potential of the personality traits approach to identify reliability related characteristics of the dataset's *headlines*. R is a free software environment for statistical computing and graphics. It compiles and runs on a wide variety of UNIX platforms, Windows and MacOS.[4] The main steps of the algorithm are:

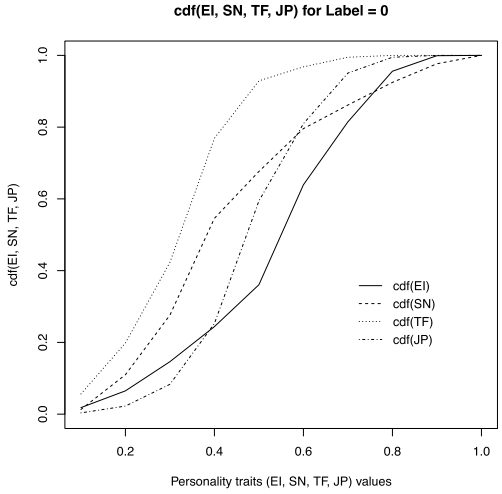

Fig. 4. Cumulative distribution function for each personality trait and Label = 0.

Step 1: the dataset is read and stored in a dataframe called `raw.comments.df`.

Step 2: the dataset is split into test and evaluation datasets, `test.comments.df` and `eval.comments.df`. The test dataset comprises 40% of the *headlines*.
The personality traits are designated as:
- Extroverts - Introverts – EI;
- Sensors - Intuitives – SN;
- Thinkers - Feelers – TF;
- Judgers - Perceivers – JP.

Step 3: the test dataset is split into two datasets, one containing the Real *headlines* (`test.comments.0.df`), i.e., the *headlines* for which Label = 0, and one containing the Fake *headlines* (`test.comments.1.df`), i.e., those headlines for which Label = 1.
The number of Real and Fake test *headlines* are 1156 and 1188, respectively.

Step 4: the cumulative distribution function (cdf) for each personality trait given the value of the label, given by Eq. (1), is evaluated.

$$\mathrm{cdf}_x(a|v) = \mathrm{Prob}(x \le a|\mathrm{label} = v) \tag{1}$$

where $x \in \{EI, SN, TF, JP\}$, $0 \le a \le 1$ and $v \in \{0, 1\}$.

Step 5: using the Bayes theorem, the probability of the label assuming a determined value given the probability of a personality trait is less than or equal to a certain value is evaluated by Eq. (2).

$$\mathrm{Prob}(\mathrm{label} = v|x \le a) = \frac{\mathrm{Prob}(x \le a|\mathrm{label} = v)}{\sum_{v\in\{0,1\}} \mathrm{Prob}(x \le a|\mathrm{label} = v)} \tag{2}$$

where $x \in \{EI, SN, TF, JP\}$, $0 \le a \le 1$ and $v \in \{0, 1\}$.

Figs. 4 and 5 show the cumulative distribution function for each personality trait, Label = 0 and Label = 1. It can be seen that the distributions are very similar indicating that it may not be possible to differentiate between the labels using the personality traits.

Figs. 6, 7, 8 and 9 show the probability of the label assuming a determined value given the probability of a personality trait (EI, SN, TF or JP) is less than or equal to a certain value.

[2] https://kaggle.com/datasets/rchitic17/real-or-fake, Accessed May 25, 2022.

[3] https://github.com/wiredtoserve/datascience/tree/master/PersonalityDetection, Accessed May 25, 2022.

[4] R can be downloaded from https://www.r-project.org.

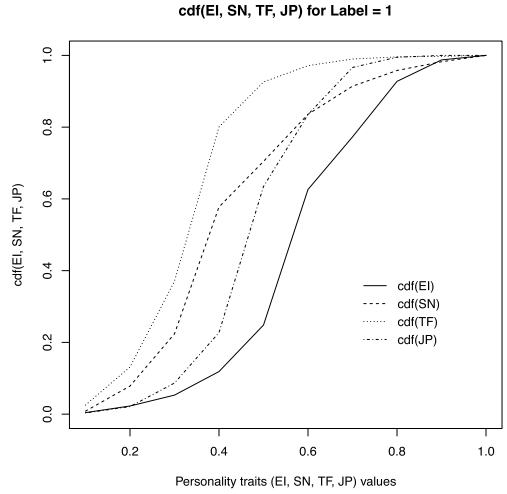

Fig. 5. Cumulative distribution function for each personality trait and Label = 1.

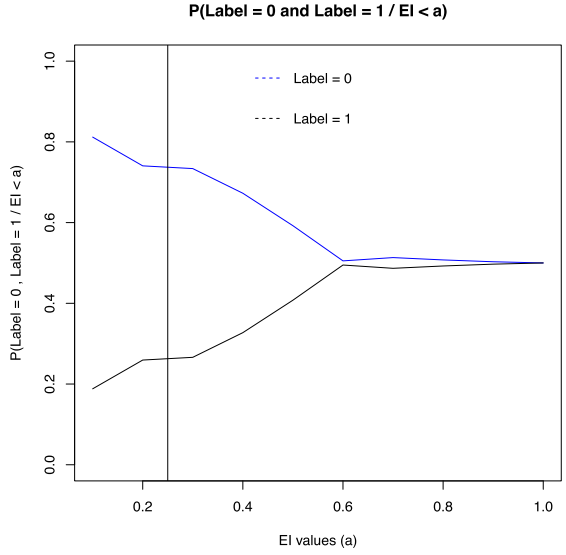

Fig. 6. Probability of the label assuming a determined value given the probability of a personality trait is less than or equal to a certain value.

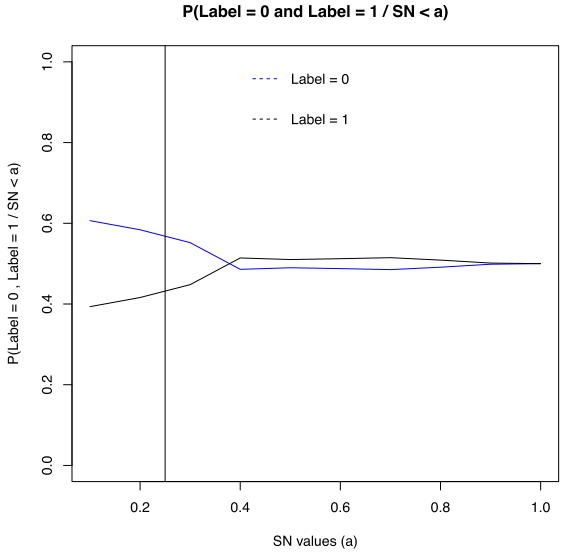

Fig. 7. Probability of the label assuming a determined value given the probability of a personality trait is less than or equal to a certain value.

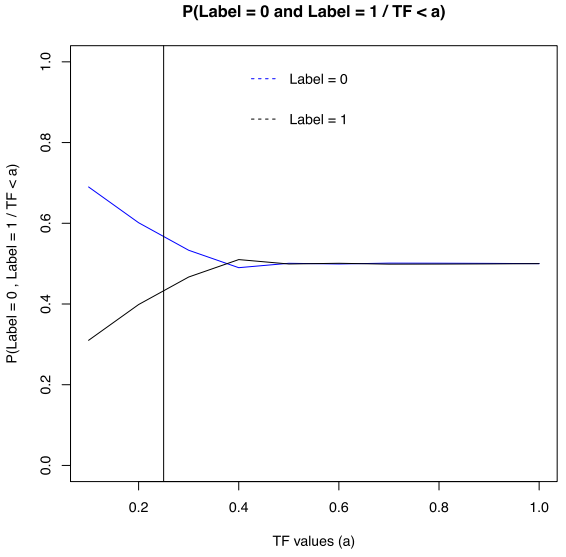

Fig. 8. Probability of the label assuming a determined value given the probability of a personality trait is less than or equal to a certain value.

It can be seen in Figs. 6, 7, 8 and 9 that: i) JP does not help at all to infer the labels values; ii) from a certain value of the EI, SN and TF personality trait it is not possible to infer the labels values. It could conjecture that low values personality traits have some differentiation power.

Step 6: a threshold limit of 0.25 was defined for the personality traits values and several simple algorithms were developed to evaluate their differentiation power. The first one is based on the total number of low value personality traits, assuming at least two traits with low values. The following three algorithms, each focuses on a single trait of low value. The final one requires that one of the traits has low value.

The obtained results are summarised as follows:

Global evaluation: total number of evaluation *headlines*: 3516.

- tested condition: if the total number of low value personality traits is greater than **1** then the predicted label is set to 0; otherwise the predicted label is set to 1;
- total number of errors: 1674;
- accuracy: 52.39%

- tested condition: if the total number of low value personality traits is greater than **2** then the predicted label is set to 0; otherwise the predicted label is set to 1;
- total number of errors: 1792;
- accuracy: 49.03%

- tested condition: if the total number of low value personality traits is greater than **3** then the predicted label is set to 0; otherwise the predicted label is set to 1;
- total number of errors: 1796;
- accuracy: 48.93%

Low EI values performance: number of *headlines* with low EI personality trait values: 55.

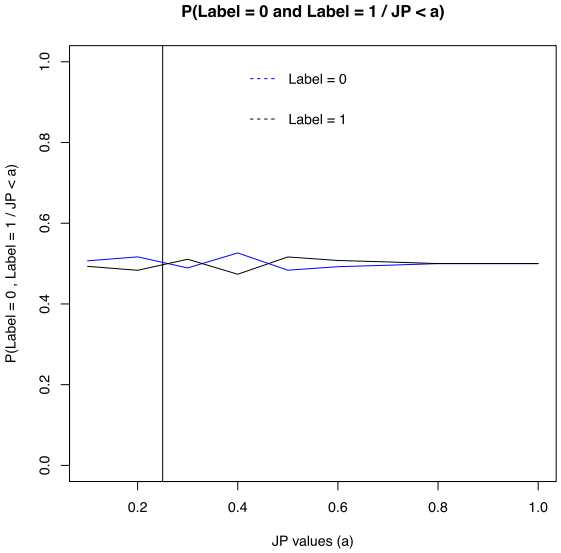

Fig. 9. Probability of the label assuming a determined value given the probability of a personality trait is less than or equal to a certain value.

- tested condition: if the *headline's* EI personality trait value is less than the threshold limit and the SN and TF values are greater than the threshold limits then the predicted label is set to 0; otherwise the predicted label is set to 1;
- number of correct predictions: 37;
- accuracy: 62.27%

Low SN values performance: number of *headlines* with low SN personality trait values: 364.

- tested condition: if the *headline's* SN personality trait value is less than the threshold limit and the EI and TF values are greater than the threshold limits then the predicted label is set to 0; otherwise the predicted label is set to 1;
- number of correct predictions: 210;
- accuracy: 57.69%

Low TF values performance: number of *headlines* with low SN personality trait values: 582.

- tested condition: if the *headline's* SN personality trait value is less than the threshold limit and the EI and TF values are greater than the threshold limits then the predicted label is set to 0; otherwise the predicted label is set to 1;
- number of correct predictions: 299;
- accuracy: 51.37%

Low personality traits values global performance: total number of *headlines* with low personality trait values: 1001.

- tested condition: if any of the last three tested conditions is satisfied then the predicted label is set to 0; otherwise the predicted label is set to 1;
- number of correct predictions: 546;
- accuracy: 54.54%

The results suggest that the use of personality traits to predict if a *headline* is Real or Fake leads to no better than a random assignment. Even the 62.27% accuracy obtained for the low EI values case cannot be considered a good result as just 55 *headlines* fall into this category.

As the results presented in this section correspond to very simple and specific criteria, the MTBI approach is further examined in Section 5.

Table 2
$\mathcal{L} = (M, A, U, H)$ assignment for Labels equal to 0 and 1.

Label = 0				Label = 1			
M	**A**	**U**	**H**	**M**	**A**	**U**	**H**
0	1	0	0	0	0	0	1
0	1	0	1	0	0	1	0
0	1	1	0	0	0	1	1
1	0	0	0	0	1	1	1
1	0	0	1	1	0	1	0
1	1	0	0	1	0	1	1
1	1	0	1	1	1	1	1
1	1	1	0				

4.3. Evaluation of the Lacanian discourses approach

In this section we examine the potential of the Lacanian discourses being used to predict Real and Fake news from their *headlines*.

At this initial stage, it has been decided to adopt a very simple way to quantify the presence of each possible Lacanian discourse in an enunciation. Consider the vector $\mathcal{L} = (M, A, U, H)$ where M, A, U, H stands for Master, Analyst, University and Hysteric, respectively, and may take the value 1 to indicate the presence of that type of discourse in the enunciation, and 0 otherwise. For example, $\mathcal{L} = (M, A, U, H) = (1, 0, 1, 0)$ indicates that in the corresponding enunciation traces of Master and University discourses have been identified.

The following procedure has been adopted to evaluate the Lacanian discourses approach:

Step 1: blind assignment of the Lacanian discourses. An expert having no access to the *headlines'* labels assign $\mathcal{L} = (M, A, U, H)$ to a set of 100 *headlines*.
Step 2: identification of ambiguities. The expert accesses the labels of the 100 used *headlines* and verifies if the same values of $\mathcal{L} = (M, A, U, H)$ have been assigned to *headlines* with different labels.
Step 3: non-blind re-assignment of the Lacanian discourses to get Zero ambiguities. For each of the identified ambiguities in Step 2, the expert verifies if it is possible to assign a **psychoanalytically valid** alternative value of $\mathcal{L} = (M, A, U, H)$ to solve the ambiguity.
Step 4: no-blind extension the Lacanian discourses assignment. Using the same criteria identified in Step 3 to solve the ambiguities the assignment of $\mathcal{L} = (M, A, U, H)$ was made for additional 200 *headlines*.

Using the aforementioned procedure it was possible to find a partition of the $\mathcal{L} = (M, A, U, H)$ codes for the 300 *headlines*. Table 2 shows the obtained partition.

It can be argued that the use of a non-blind reassignment of $\mathcal{L} = (M, A, U, H)$ is not fair and is just trick to get zero ambiguity. It is an understandable argument, however the procedure is justified by the following reasons:

1. Psychoanalysis is not a hard science and different experts adopting different points of view may arrive at different valid conclusions. The important point to be emphasised is that a psychoanalytically valid partition of the $\mathcal{L} = (M, A, U, H)$ codes have been found.
2. The length of the *headlines* is very short, much shorter than a normal patient's narrative. In this case the assignment of $\mathcal{L} = (M, A, U, H)$ is much harder and a second turn of re-assignment is mandatory to arrive at useful results.
3. The assignment of $\mathcal{L} = (M, A, U, H)$ to *headlines* is quite different from the identification of the discourse adopted by the patient in a psychoanalytical setting where besides the narrative, body language, clothes, emotion representative gestures help the analyst to build a full representation of the patient.

To illustrate the difficulty of assigning $\mathcal{L} = (M, A, U, H)$ consider the following *headline*:

Police Turn In Badges Rather Than Incite Violence Against Standing Rock Protestors

to which has been assigned $\mathcal{L} = (M, A, U, H) = (1, 0, 1, 0)$.

The *Master* assignment is justified by the mention of "Police" that is a figure of authority. The *University* assignment is justified because this *headline* is disclosing an information, some kind of knowledge. However, it could be argued that a $\mathcal{L} = (M, A, U, H) = (1, 0, 1, 1)$ assignment is better because the reference to "Violence" is a trace of a *Hysteric* discourse. Is this assignment was adopted nothing would change because both codes are in the Label $= 1$ partition. On the other hand, it could also be argued that a $\mathcal{L} = (M, A, U, H) = (1, 1, 0, 0)$ is better because the verb "Incite" may be taken more as an opinion and interpretation than a mere disclosure of information. If this assignment were adopted, the new code belongs to the Label $= 0$ partition and an ambiguity would have been introduced.

This simple example shows that the assignment of Lacanian discourses is not an easy task and requires further investigation.

From Table 2 a deterministic model to predict the labels can be easily derived.

The Label should be set to Zero if $\mathcal{L} = (M, A, U, H)$ satisfies Eq. (3), i.e., the substitution of $\mathcal{L} = (M, A, U, H)$ into Eq. (3) results in 1.

$$A.\overline{U} + M.\overline{U} + A.U.\overline{H} \tag{3}$$

The Label should be set to One if $\mathcal{L} = (M, A, U, H)$ satisfies Eq. (4), i.e., the substitution of $\mathcal{L} = (M, A, U, H)$ into Eq. (4) results in 1.

$$\overline{A}.U + \overline{M}.\overline{A}.H + U.H \tag{4}$$

Equations (3) and (4) are minimum, but are not unique. They complement each other in the sense that if the substitution of the $\mathcal{L} = (M, A, U, H)$ values satisfy (3) they will not satisfy (4), and vice-versa.

Table 2 and Equations (3) and (4) suggest that the presence of the Analyst discourse, $\mathcal{L} = (M, A, U, H) = (*, 1, *, *)$ almost always means that the corresponding *headline* refers to a Real news. On the other hand, the presence of the University discourse, $\mathcal{L} = (M, A, U, H) = (*, *, 1, *)$ almost always means that the corresponding *headline* refers to a Fake news.

Finally the assignment of $\mathcal{L} = (M, A, U, H)$ values was extended to another 300 *headlines* to allow further evaluation of the procedure in Section 5.

5. Computational approach

This section describes the Machine Learning (ML) experiments of our approach to classify and predict whether a *headline* is Fake or Real. The following subsections describe the various adopted procedures and showcase the results.

5.1. Algorithmic personality type (MTBI) assignment to the dataset's headlines

Several open-source codes and projects have been developed to extract the personality types from a text as, for example, the works available in Nagpurkar (2020) and Saini and Agarwal (2020).

In this work, the dataset file was submitted to the personality traits assignment algorithm developed by Nagpurkar (2020), which is implemented in the programming language Python.[5] The analysis is performed using linguistic cues like word repetition, number of verbs, nouns, detection of emotion in text and so on. We are mainly interested in the application of such a theory and not in the implementation of the theory itself as it has already been done and evaluated by other researchers as in Iskandar et al. (2021). Some of the approaches are able to determine types like E/I, S/N and T/F with at least 90% accuracy Li et al. (2018) and, more recently, the J/P type with 81% and 65% accuracy, depending on the used dataset Choong and Varathan (2021).

The MBTI algorithm takes as input a text and gives as output 4 groups of variables (Extroversion/Introversion, Thinking/Feeling and so on) in the range of [0,1], which represent the dimensions described in Section 3.1, and the binary classification task of the ML models is to classify the *headlines* whether they are Real (dataset label is 0) or Fake (dataset label is 1).

The mean of each personality type group is presented in Table 3.

The distribution of the obtained types is illustrated in Figs. 10 (Thinking - Feeling), 11 (Introversion - Extroversion), 12 (Judging - Perceiving) and 13 (Sensing - Intuiting).

Table 3
Mean, rounded to 2 decimals, of each variable (type) obtained from the algorithm.

Personality Type	Mean
Introversion - Extroversion	0.45 & 0.55
Sensing - Intuiting	0.43 & 0.57
Thinking - Feeling	0.33 & 0.67
Judging - Perceiving	0.49 & 0.51

Table 4
The summary of the performance of the ML algorithms and the relative achieved balanced accuracy.

Model	Type 1	Type 2	Type 3
Extra-Trees-Classifier	58%	57%	56%
LightGBM	59%	59%	58%
SVC	57%	57%	55%
GaussianNB	54%	54%	55%

5.2. MTBI fake or real news prediction using ML

We selected four different ML algorithms, namely: i) Support Vector Classifier (SVC), ii) LightGBM, iii) Extra-Trees-Classifier and iv) GaussianNB (GNB) for comparison and tested them with the following groups of independent personality types in an attempt to identify the impact of different personality types in the field of Fake/Real news detection:

- Type 1: Introversion, Intuition, Thinking, Judging (INTJ)
- Type 2: Extroversion, Sensing, Feeling, Perceiving (ESFP)
- Type 3: Introversion, Sensing, Thinking, Perceiving (ISTP)

Each algorithm is run using default parameters, is trained with 25% of data, and the results are presented in Table 4. It can be seen that Types 1, 2 led to 59% accuracy by the LightGBM algorithm, while Type 3 had almost always the lowest score. The GaussianNB had the worst performance among all. It is noticeable that hyperparameter tuning didn't improve the accuracy and didn't provide any additional insight about the characteristics of the task and dataset.

The confusion matrix derived from the LightGB algorithm, Type 1 experiment, is shown in Fig. 14. It can be seen that the best-scoring model is able to correctly predict a fake news *headline* as Fake 30.2% of the time (bottom right corner of figure), while there is a probability of approximately 22% to categorize a fake news *headline* as Real (bottom left). Regarding the real news label, approximately 27.9% correctly categorize a real news *headline* as Real (top left) and 19.8% correctly to categorize a real news headlines as Fake (top right).

[5] https://www.python.org/.

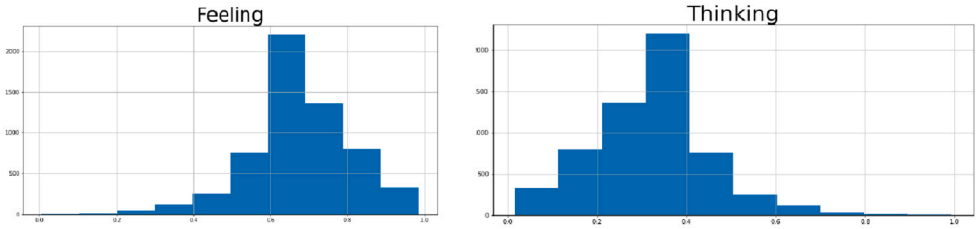

Fig. 10. Distribution of types Feeling (mean 0.67) and Thinking (mean 0.33).

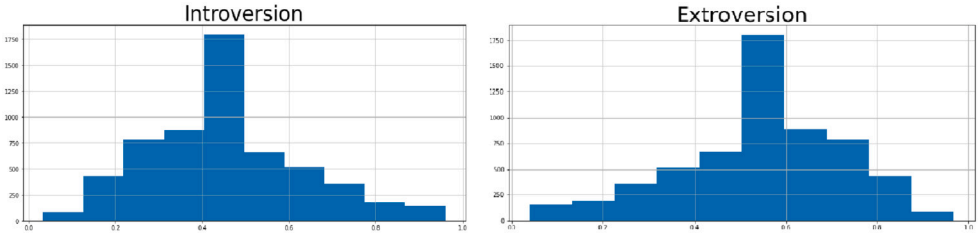

Fig. 11. Distribution of types Introversion (mean 0.49) and Extroversion (mean 0.55).

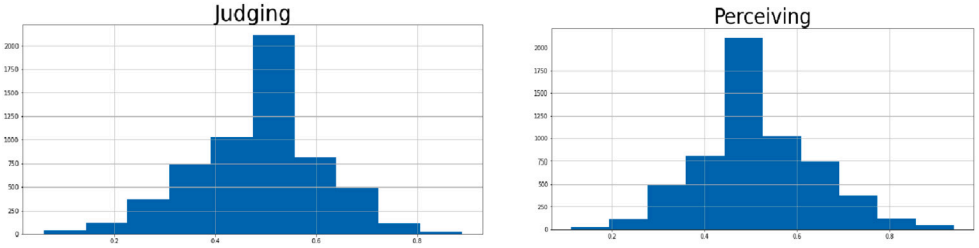

Fig. 12. Distribution of types Judging (mean 0.45) and Perceiving (mean 0.51).

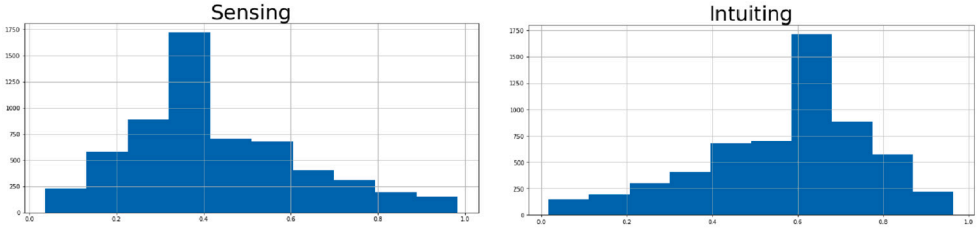

Fig. 13. Distribution of types Sensing (mean 0.43) and Intuiting (mean 0.57).

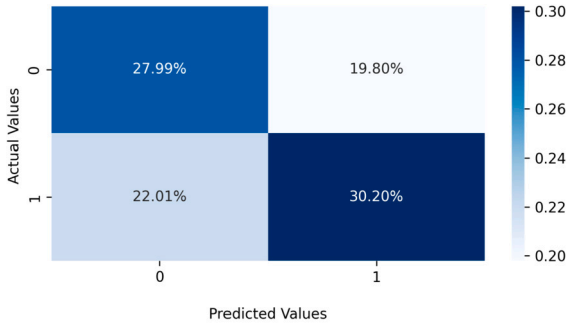

Fig. 14. Confusion Matrix of the LightGBM algorithm for the Type 1 experiment.

The obtained results confirm what has already been stated in Section 4.2 that MBTI is not the ideal tool for predicting *headline* labels, as none of the ML algorithms achieved an accuracy score much better than a random prediction.

5.3. Lacanian discourses fake or real news prediction using ML

This is a two steps procedure:

- Step 1: Assign the Lacanian Discourses to the *headlines* either manually or automatically.
- Step 2: Using the results from Step 1, predict the labels of the *headlines*, as Fake or Real.

This gives rise to four possible combinations of discourses assignment and labels prediction:

- *Manually* assign the Lacanian Discourses and predict labels using the deterministic model discussed in Section 4.3 (Section 5.3.1).
- *Manually* assign the Lacanian Discourses and predict labels using ML algorithms (Section 5.3.2).
- *Automatically* assign the Lacanian Discourses using the language model GPT-3 and predict labels using the deterministic model discussed in section 4.3 (Section 5.3.3).
- *Automatically* assign the Lacanian Discourses using the language model GPT-3 and predict labels using ML algorithms (Section 5.3.4).

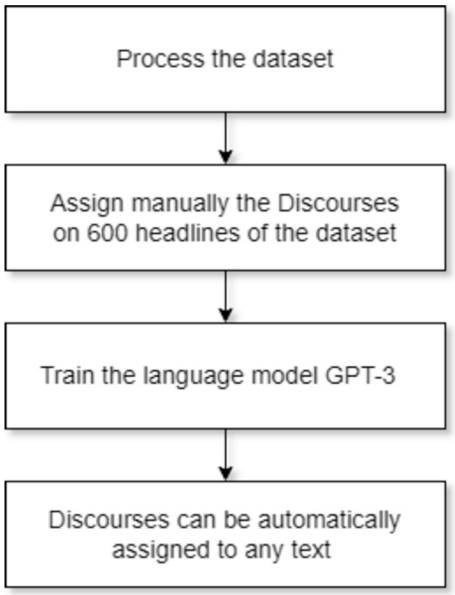

Fig. 15. The 4 steps to achieve the automatic assignment of Lacanian Discourses. After the processing the dataset, we assign manually the discourses to 600 random headlines of the dataset, then we train the language model GPT-3 resulting in the last step to be able to assign automatically the Lacanian discourses to the rest of headlines of the dataset.

For the first two approaches, we randomly selected 600 *headlines* and manually assigned the corresponding Lacanian discourses.

In the case of the latter two approaches, as shown in Fig. 15, we used the aforementioned 600 manually assigned *headlines* to train the language model GPT-3 (Generative Pre-trained Transformer 3) in an attempt to achieve automatic Lacanian Discourses assignment to solve the problem of *Step 1*.

GPT-3 is a neural network machine learning model, developed by OpenAI, trained using internet data to generate any type of text that enables developers to train and deploy AI models. It provides a variety of tools and services for data preprocessing and model training, and its capabilities include but are not limited to: i) generating, classifying, translating and summarising text; ii) generating and answering questions; iii) generating images and audio from text Dale (2020) and can be accessed and used on OpeanAI's main website.[6]

5.3.1. Manual Lacanian discourses assignment and deterministic model labels prediction (no ML used)

This case was introduced and implemented in Section 4.3. The accuracy is almost 100%, however, it is not a practical alternative because the manual assignment of Lacanian discourse is only feasible for a very small number of *headlines*. Nevertheless, it is an important theoretical case because it has shown the existence of a psychoanalytic valid partition of the codes.

5.3.2. Manual Lacanian discourses assignment and ML labels prediction

The ML algorithms presented in Section 5.2 were trained and evaluated with the 600 manually Lacanian discourses assigned to the *headlines*. The accuracy of the labels prediction is summarised in Table 5.

As expected, the accuracy is slightly decreased compared to the first approach. It is important to realise that this is not a typical ML scenario as the dataset is very small and the models may be overfitting: the model fits perfectly against its training data and thus achieves very high accuracy.

Table 5
Performance of the ML algorithms and the relative achieved balanced accuracy of labels prediction for manually assigned Lacanian discourses.

Model	Accuracy
Extra-Trees-Classifier	97%
LightGBM	97%
SVM	97%
GaussianNB	92%

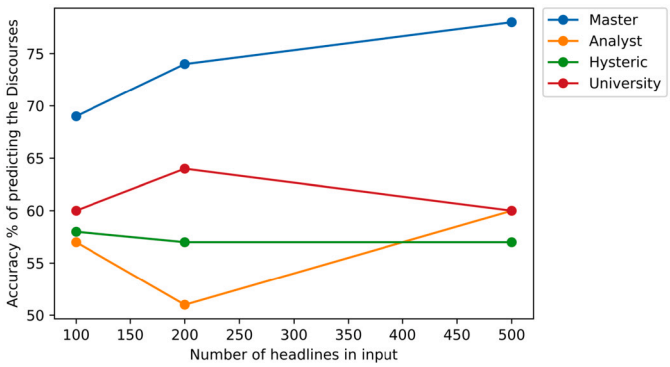

Fig. 16. The X axis represents the number of manually assigned *headlines* used to train the classification model; the Y axis represents the accuracy of predicting the discourses of the last 100 manually assigned *headlines*.

Once again we emphasise that the manual assignment of Lacanian discourses is rather impractical but it confirms the potential of the psychoanalytic-based method.

5.3.3. Automatic Lacanian discourses assignment and deterministic model labels prediction

OpenAI[7] provides an API to work with four different models: Davinci, Curie, Babbage, and Ada. Each one of them has different capabilities, advantages and disadvantages, and after some experimenting we decided to work with the model Ada because of its speed and its capabilities more adapted to the nature of the classification under study.

We first created a set of classification Ada models using 100, 200, and 500 out of 600 manually Lacanian discourses assigned *headlines* to evaluate its accuracy in predicting the discourses of the last 100 *headlines*.

The results of this experiment are shown in Fig. 16. It can be seen that as the model is getting trained with more assigned *headlines*, it gets better at classifying and labelling the Master discourse, while it gets worse at labelling the Analyst discourse and a bit better at labelling the University discourse when it is trained with 200 headlines, but it ends up at the same accuracy of 100 assigned *headlines* when the input is 500 assigned *headlines*.

The overall accuracy of exactly predicting the four discourses of the last 100 *headlines* is illustrated in Fig. 17. It can be seen that the accuracy is below 30% but with a slight increase as the number of assigned discourses increases.

Giving the 600 manually assigned *headlines* as training in the Ada model and evaluating the whole dataset using the deterministic model described in Section 4.3, the labels prediction accuracy is only 66%. This is a poor result, however much better than what was obtained with the MTBI approach, due to a poor automatic discourses assignment accuracy, as discussed before.

[6] https://openai.com/api/.

[7] https://beta.openai.com/docs/models/gpt-3.

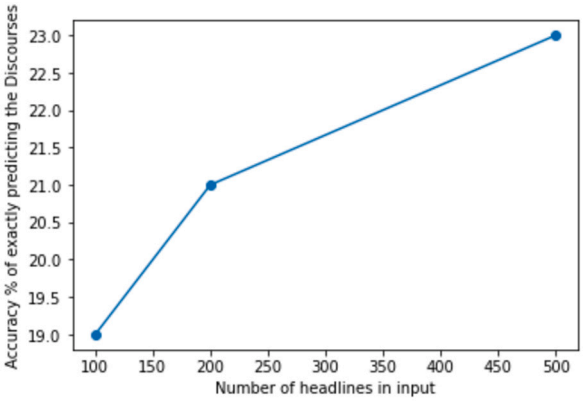

Fig. 17. The X axis represents the number of manually assigned *headlines* used to train the classification model; the Y axis represents the accuracy of exactly predicting the four discourses of the last 100 manually assigned *headlines*.

Table 6
The ML models' performance and the relative achieved balanced labels prediction accuracy.

Model	Number of training points		
	100	300	600
Extra-Trees-Classifier	63%	64%	66%
LightGBM	63%	64%	66%
SVC	63%	64%	66%
GaussianNB	62%	63%	65%

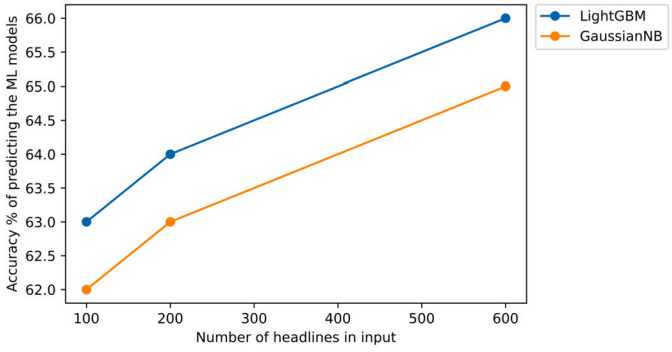

Fig. 18. The X axis represents the number of manually assigned *headlines* used to train the classification model; the Y axis represents the ML models dataset labels prediction accuracy.

5.3.4. Automatic Lacanian discourses assignment and ML labels prediction

Following the same procedure as described in Section 5.3.3, using 100, 300, and 600 as training points, the results of the ML models labels prediction of the *headlines* are shown in Table 6. The Extra-Trees-Classifier, LightGBM, SVC, and GaussianNB models best accuracy is 66%.

The accuracy gets better as the number of training points increases, as shown in Fig. 18.

To estimate how many headlines should be used for training the models to achieve a 70% accuracy threshold we did a log-fitting of the obtained points using Eq. (5):

$$a\log(bx+1)+0.5. \tag{5}$$

Fig. 19, that is in logarithmic scale, shows the results.

The parameters for the log function are:

- GNB: a = 0.0198352, b = 3.56483;
- SVC: a = 0.0130760, b = 312.088.

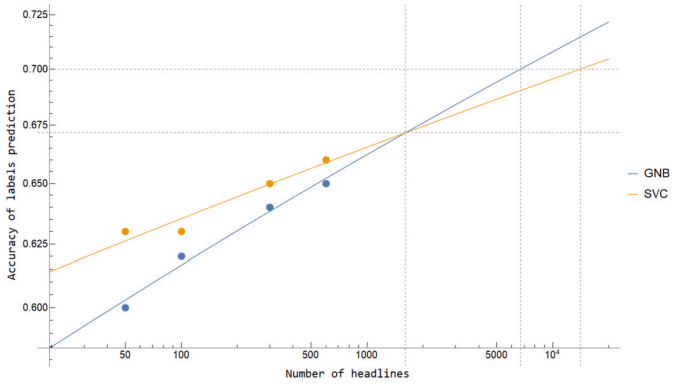

Fig. 19. The X axis represents the number of manually assigned *headlines* used to train the classification model; the Y axis the dataset labels prediction accuracy of the GNB (blue), SVC (yellow) ML models.

The number of points needed to a 70% accuracy threshold is 6700 for the GNB model, and 14000 for the SVC model. With 1600 points the two models provide the same 67% accuracy, and after that point the GNB model performs better.

Despite the power of models provided by the OpenAI the problem of automatically assign the Lacanian discourses remains to be solved to benefit from the great potential of psychoanalytic-based approach to be used to extract information from enunciations.

6. Discussions

Compared to all psychological-based approaches, the approach proposed in this work is fundamentally different since we adopt a psychoanalytic perspective; in particular, we employ the powerful notion of Lacanian discourse types.

The evaluation of the psychological-based approach of deriving the Real or Fake characteristic from the personality traits associated to each *headline*, done in Section 4, has shown that this approach is not better than a random prediction. This conclusion has been confirmed in Section 5.2, where different ML algorithms achieved at most 56% accuracy in label prediction.

The new psychoanalytic-based is a two steps procedure: i) each *headline* received a Lacanian discourses assignment $\mathcal{L}=(M,A,U,H)$, where M, A, U, H stands for Master, Analyst, University and Hysteric, respectively, and may take the value 1 to indicate the presence of that type of discourse in the enunciation, and 0 otherwise.; ii) Real or Fake label prediction from the $\mathcal{L}=(M,A,U,H)$.

A partition of the $\mathcal{L}=(M,A,U,H)$ between Real and Fake labels have been obtained and a deterministic model for label prediction has been derived. This indicates the possibility of achieving a **theoretical 100% accuracy of label prediction**. However, the described procedure is not practical as the Lacanian discourses assignment was manually done. The results were confirmed in Section 5.3.

A first attempt to automatically assign the Lacanian discourses using the environment provided by OpenAI resulted in a label prediction accuracy decrease to 66% but still significantly higher than the psychological-based approach.

7. Conclusions

Different approaches of combining psychological and social dimensions with computational methods which include computational psycholinguistics, personality traits, behavioural analysis, emotional states and cognitive psychology methods have been used to derive information from user-related and generated data. However, in various contexts where user perception is more susceptible to emotional or some other form of bias, the reliability of user data is often compromised.

Compared to all those approaches, the approach proposed in this work is fundamentally different since we adopt a psychoanalytic perspective based on the Lacanian discourse types concept. To the best of our knowledge, this is the first attempt to systematically bring together psychoanalysis and computing.

From the experimental point of view, as a preliminary case study scenario, we looked at the problem of identifying from the *headlines* of published news if they are Real or Fake.

In Section 4, it has been shown that the approach based on the notion of personality traits is not better than a random prediction. Different ML algorithms achieved at most 56% accuracy in label prediction.

On the other hand, for the psychoanalytic-based approach a partition of the $\mathcal{L} = (M, A, U, H)$ of manually assigned discourses between Real and Fake labels have been obtained and a deterministic model for label prediction has been derived. This indicates the possibility of achieving a theoretical 100% accuracy of label prediction. Automatic assignment of the Lacanian discourses was done using the environment provided by OpenAI. The label prediction accuracy decreased to 66%.

In summary, we conclude that the psychoanalytic-based approach has a great potential to extract information from user-generated data.

As future work, an automatic procedure to assign the Lacanian discourses will be developed, and the technique will be applied to other problems rather than the Real / Fake news detection. The procedure will be extended to non-binary classification problems and to multimedia user-generated data.

CRediT authorship contribution statement

Minas Gadalla: Methodology, Data gathering, ML research and simulation, paper writing administration. **Sotiris Nikoletseas:** Validation, Formal analysis, Writing - Review & Editing. **Jose Roberto de A. Amazonas:** Conceptualization, Model development and validation, Writing - Review & Editing. **Jose D.P. Rolim:** Original proposal coordination, critical review.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgement

This publication is part of the Spanish I+D+i project TRAINER-A (ref. PID2020-118011GB-C21), funded by MCIN/AEI/10.13039/501100011033.

References

Bailly, L. (2009). *Lacan: A beginner's guide*. Oneworld Publications.

Boyle, G. J. (1995). Myers-briggs type indicator (MBTI): Some psychometric limitations. *Australian Psychologist*, *30*, 71–74. https://doi.org/10.1111/j.1742-9544.1995.tb01750.x.

Cardaioli, M., Cecconello, S., Conti, M., Pajola, L., & Turrin, F. (2020). Fake news spreaders profiling through behavioural analysis. In *CLEF 2020 proceedings* (pp. 1–9).

Castro, J. C. L. d. (2016). Applications of the lacanian theory of discourses in the field of communication. *Intercom, Rev. Bras. Ciênc. Commun. [online]*, 99–113.

Cattell, R. B. (1943). The description of personality: basic traits resolved into clusters. *Journal of Abnormal and Social Psychology*, *38*, 476–506. https://doi.org/10.1037/h0054116.

Chen, T.-Y., Chen, Y.-M., & Tsai, M.-C. (2021). A status property classifier of social media user's personality for customer-oriented intelligent marketing systems. In *Research anthology on strategies for using social media as a service and tool in business* (pp. 557–581). IGI Global.

Choong, E. J., & Varathan, K. D. (2021). Predicting judging-perceiving of myers-briggs type indicator (MBTI) in online social forum. *PeerJ Computer Science*, *9*, Article e11382. https://doi.org/10.7717/peerj.11382.

Ciampaglia, G., Shiralkar, P., Rocha, L., Bollen, J., Menczer, F., & Flammini, A. (2015). Computational fact checking from knowledge networks. *PLoS Computational Biology*, *10*. https://doi.org/10.1371/journal.pone.0128193.

Clark, L., & Çallı, L. (2014). Personality types and facebook advertising: An exploratory study. *Journal of Direct Data and Digital Marketing Practice*, *15*, 327–336. https://doi.org/10.1057/dddmp.2014.25.

Conroy, N., Rubin, V., & Chen, Y. (2015). Automatic deception detection: Methods for finding fake news. In *Proceedings of the Association for Information Science and Technology* (pp. 1–4).

Contri, G. B. (Ed.). (1978). *Lacan in Italia 1953-1978* (pp. 32–55). Milano: La Salamandra.

Dale, R. (2020). GPT-3: What's it good for? *Natural Language Engineering*, *27*, 113–118. https://doi.org/10.1017/s1351324920000601.

Feng, V., & Hirst, G. (2013). Detecting deceptive opinions with profile compatibility. In *Proceedings of the sixth international joint conference on natural language processing* (pp. 338–346).

Freud, S. (1914). The moses of michelangelo. In J. Strachey (Ed.), *The standard edition of the complete psychological works of Sigmund Freud, volume XIV* (pp. 211–238). London: Hogarth Press.

Freud, S. (1921). Group psychology and the analysis of the ego. In J. Strachey (Ed.), *The standard edition of the complete psychological works of Sigmund Freud, volume XVIII* (pp. 67–143). London: Hogarth Press.

Freud, S. (1927). The future of an illusion. In J. Strachey (Ed.), *The standard edition of the complete psychological works of Sigmund Freud, volume XXI*. London: Hogarth Press (pp. 3–56).

Freud, S. (1930). Civilization and its discontents. In J. Strachey (Ed.), *The standard edition of the complete psychological works of Sigmund Freud, volume XXI*. London: Hogarth Press (pp. 59–145).

Freud, S. (1933). Why war? In J. Strachey (Ed.), *The standard edition of the complete psychological works of Sigmund Freud, volume XXII* (pp. 197–215). London: Hogarth Press.

Guo, X., Sun, Y., & Vosoughi, S. (2021). Emotion-based modeling of mental disorders on social media. In *IEEE/WIC/ACM international conference on web intelligence (WI-IAT '21)* (pp. 8–16).

Iskandar, A. F., Utami, E., & Hidayat, W. (2021). A web-based indonesian MBTI prediction: Design, implementation, and testing. In *2021 IEEE 5th international conference on information technology, information systems and electrical engineering (ICITISEE)* (pp. 169–174). IEEE.

John, O. P., & Robins, R. W. (1994). Traits and types, dynamics and development: No doors should be closed in the study of personality. *Psychological Inquiry*, *5*, 137–142. https://doi.org/10.1207/s15327965pli0502_10.

Jung, C. G. (1974). *Psychological types*. Princeton University Press.

Kachur, A., Osin, E., Davydov, D., Shutilov, K., & Novokshonov, A. (2020). Assessing the big five personality traits using real-life static facial images. *Scientific Reports*, *10*. https://doi.org/10.1038/s41598-020-65358-6.

Kaplan, R. M. (2019). Computational psycholinguistics. *Computational Linguistics*, *45*, 607–626.

Koury, A., & Hernandez, R. (2019). NLP for Fake News Classification. Available in https://www.kaggle.com/code/akoury/nlp-for-fake-news-classification.

Kumar, K. P. K., & Geethakumari, G. (2014). Detecting misinformation in online social networks using cognitive psychology. *Human-Centric Computing and Information Sciences*, 4–14.

Lacan, J. (1972). Seminaire 20: Encore. Available in http://staferla.free.fr/S20/S20%20ENCORE.pdf.

Lacan, J. (1974). Seminaire 22: R.S.I. Available in http://staferla.free.fr/S22/S22%20R.S.I..pdf.

Li, C., Hancock, M., Bowles, B., Hancock, O., Perg, L., Brown, P., Burrell, A., Frank, G., Stiers, F., Marshall, S., Mercado, G., Ahmed, A.-W., Beckelheimer, P., Williamson, S., & Wade, R. (2018). Feature extraction from social media posts for psychometric typing of participants. In *Augmented cognition: Intelligent technologies* (pp. 267–286). Springer International Publishing.

Murray, H. A., & McAdams, D. P. (2007). *Explorations in personality*. Oxford University Press.

Nagpurkar, M. (2020). Personality detection: NLP. Available in https://github.com/wiredtoserve/datascience/tree/master/PersonalityDetection. (Accessed 25 May 2022).

Rothmann, S., & Coetzer, E. P. (2003). The big five personality dimensions and job performance. *Journal of Industrial Psychology*, *29*. https://doi.org/10.4102/sajip.v29i1.88.

Saini, S., & Agarwal, P. (2020). Classification of personality based on users Twitter data. Available in https://github.com/priyansh19/Classification-of-Personality-based-on-Users-Twitter-Data. (Accessed 25 May 2022).

Sampat, B., & Raj, S. (2016). Fake or real news? understanding the gratifications and personality traits of individuals sharing fake news on social media platforms. *Aslib Journal of Information Management*.

Sastrawan, I. K., Bayupati, I., & Arsa, D. M. S. (2022). Detection of fake news using deep learning CNN–RNN based methods. *ICT Express*, *8*, 396–408. https://doi.org/10.1016/j.icte.2021.10.003.

Smith, C. P. (Ed.). (1992). *Motivation and personality*. Cambridge University Press.