International Journal of Electrical and Computer Engineering (IJECE)
Vol. 13, No. 6, December 2023, pp. 6984~6991
ISSN: 2088-8708, DOI: 10.11591/ijece.v13i6.pp6984-6991

Approach for Enneagram personality detection for Twitter text: a case study

Esraa Abdelhamid, Sally S. Ismail, Mostafa Aref
Computer Science Department, Faculty of Computer and Information Science, Ain Shams University, Cairo, Egypt

Article Info

Article history:

Received Mar 6, 2023
Revised May 3, 2023
Accepted Jun 4, 2023

Keywords:

Enneagram
Personality detection
Personality model
Sentiment analysis
Text mining

ABSTRACT

Understanding people's emotions and orientations attracts researchers nowadays. Current personality detection research concentrates on models such as the big five model, the three-factor model. The Enneagram is deeper than these models for providing a comprehensive view. This theory is a unique personality model because it illustrates what drives human behavior. This recognition helps in building smarter recommendation systems and intelligent educational systems. Enneagram personalities are realized through a long questionnaire-based test. People are not concerned about doing a test because it is time-consuming. A proposed case study employs Twitter's text to detect Enneagram personality because it requires no time or effort. The proposed case study is based on an approach that uses a combination of ontology, lexicon, and statistical technique. This proposed case study uses the biography description text and 40 tweets of a Twitter profile text. The highest probability percentage is peacemaker personality which is 15.58%. This result means that the identified personality is the peacemaker. The outcome is equivalent to the determination of the Enneagram's specialized people. This result promises more positive outcomes. This is the first automated approach to determine the Enneagram from text.

Corresponding Author:

Esraa Abdelhamid
Computer Science Department, Faculty of Computer and Information Science, Ain Shams University
Abbasia, Cairo, Egypt
Email: esraa_cs@hotmail.com

1. INTRODUCTION

Personality detection attracts researchers nowadays. The availability of huge volumes of data in fields like psychology, text mining, machine learning, and natural language processing motivates academics to conduct studies to discover people's sentiments [1]. Identifying the personality can benefit many domains. Recommendation systems are greatly helpful by understanding people's behavior and preferences because it enables us to greatly enhance it [2]. Personality helps recognize the person's directions, likes, and preferences on social platforms. This knowledge aids in attracting more users and bringing in extra advertisements. The huge quantity of text created by users has motivated the development of author profiling [3]. Personality identification aids in selecting the right candidate in the job hiring process. This awareness can give an insight into the person's abilities which is compared to the job qualifications.

The personality assessment test is a method to decide the personality. Humans have no enthusiasm for doing a test because it is time-consuming. People are unwilling to complete a questionnaire since it is inconvenient and time-consuming [4]. There are other disadvantages to the assessment. The main concern when conducting a personality test is that responders prefer to answer in a faking manner with artificial responses [5]. Identifying the personality is a complex mission because it demands the knowledge of the specialized. Natural language understanding makes the task more complex. Personality recognition is gaining popularity in a variety of contexts including social media, robotics, speech processing, and human computer interaction [6]. Social media is playing a crucial role nowadays. Digital footprints can indicate personality traits. It can also be used as a quick method of survey with lower fees and larger population support [7]. There are other benefits of social media. Crowd behavior analysis utilizes social media users' information [8]. Personality defines the attributes of the person's characteristics. There are many personality models: the Myers-Briggs type indicator (MBTI), the three-factor model, the big five model, and Enneagram.

The Enneagram is a personality theory that understands human behavior at a depth level. The Enneagram is a mapping of human nature's nine personality types and their mutual relationships [9]. This personality model consists of nine types: reformer, helper, achiever, individualist, investigator, loyalist, enthusiast, challenger, and peacemaker. This personality model demonstrates fears, desires, features, motivations, and problems for the nine personalities. The main advantage of the Enneagram is describing what drives human behavior. It can teach a person about his/her strengths, weaknesses, and aim for his/her growth [10]. This personality model is more personal than other models. Many universities in the United States are studying the Enneagram in psychology, education, medicine, business, and arts [11].

There are many other benefits. It is a useful tool for improving friends, family, and workplace relationships [12]. It also helps the counselor. It assists the counselor in detecting client behavior which enhances development and recovery [13]. Psychological assistance is vital in helping a patient's recovery. Since 1970, psychiatrists have utilized the Enneagram [14]. Knowing this model as fast as possible aids in recovery in less time. In education, a useful tool for overall student development is the Enneagram [15]. Personality detection by the analysis of social platforms is a recent research domain in machine learning [16]. There are multiple drawbacks of machine learning in consistency, transparency, and dependability. These difficulties are especially important in the natural language processing area which prohibits artificial intelligence from obtaining human-like performance [17].

The Enneagram personality detection approach consists of four main stages: text preprocessing, feature extraction, feature selection, and personality detection. Text preprocessing purpose is to prepare the text and remove unimportant information. Feature extraction transforms preprocessed text into word-based features. Feature selection is to pick the personalities' relevant features. This stage is performed using the saurus English lexicon [18] and Enneagram ontology [19], [20] The personality detection phase uses a statistical technique to calculate the personality's probability distribution for each one. The largest probability percentage is the detected personality [21]. The approach uses Twitter Text. Twitter is a large platform that provides a huge quantity of text. Twitter is a great chance to study naturally the people's attitudes [22]. The proposed case study is applied to the public profile Twitter account of Morgan Freeman [23]. Probability distributions for each personality are calculated and resulted in 15.58%, 12.12%, 11.90%, 11.69%, 11.47%, 11.04%, 9.96%, 9.52%, and 6.71%. The largest percentage is 15.58% among personalities that is the peacemaker. The result analysis is that the peacemaker is the personality of this profile. This conclusion is identical to the Enneagram expert's analysis [24].

This paper contains several sections: the related work, the method, the proposed personality detection case study, the result, and the conclusion. The related work demonstrates recent work done on personality detection. The method describes the approach for Enneagram personality. The proposed personality detection case study demonstrates the details of the proposed case study. The result section demonstrates the result and the case study positive outcomes. The conclusion section summarizes the approach, case study, and the result inference with future direction.

2. RELATED WORK

For detecting personality from text, two strategies have been used: a machine learning technique and text linguistic properties method [25]. Recent research has used different machine learning techniques to identify personality. Different machine learning techniques are employed like unsupervised estimation and deep learning. The big five personality model is utilized in the major of the systems. The future directions are enhancing performance by using other techniques, more features, advanced language processing, testing on various platforms, more efficient parsing, and better preprocessing techniques as shown in Table 1.

There are several approaches in text for extracting personality characteristics. Several methods employ a closed-vocabulary technique having psycholinguistic tools, while others employ an open-vocabulary one. Most personality prediction research requires a dataset to execute supervised learning and obtaining social media users' personality traits dataset has a high cost. Present personality prediction can be improved in many directions like using more appropriate algorithms or preprocessing approaches to obtain greater accuracy and implementing additional personality models [26]. Artificial recognition of personality from text is a hard mission. This domain research is still challenging and needs a lot of enhancements.

Table 1. Related work

Description	Technique	Personality model	Future direction
Personality detection system was presented that integrates questionnaire-based and text-based techniques [27].	Unsupervised estimation	Big five personality model	• Develop more advanced natural language processing models. • Evaluate other systems different from personality.
Two attention architectures for incorporating emoji and textual information in personality identification were presented [28].	Deep learning architecture	Big five personality model	• Incorporate visual features to improve performance. • Test models on more social datasets. • Use other learning models to improve performance.
Personality identification system was proposed. A dictionary system is developed [29]–[31].	Ontology and five algorithms (sentence processing, database saving, csv processing, radix tree conversion model, and trait estimation)	Big five personality model	• Use a better parsing algorithm to get more accurate results. • Add more words to the platform's corpus from various social media platforms. • Give weights to classified words instead of the word's frequency only.

3. METHOD

The applied language is the English language. Google Colab is used as a development platform. The used programming language is Python. There are many applied libraries: text preprocessing python package, TextBlob, Keras text tokenizer, OwlReady2 and Tweepy. Text preprocessing step packages are text preprocessing package, TextBlob and Keras text tokenizer. Ontology access is utilized by the OwlReady2 package. Tweepy is applied to access Twitter. The experiment is performed on 16.0 GB RAM, 64-bit system type and Intel(R) Core (TM) i7-1065G7 CPU.

Tweets are selected as input to the personality detection approach. The approach utilizes both the Enneagram ontology [19], [20] and the English Lexicon [18]. Enneagram ontology represents Enneagram knowledge. The English lexicon enhances the vocabulary seeds. The general approach is shown in Figure 1.

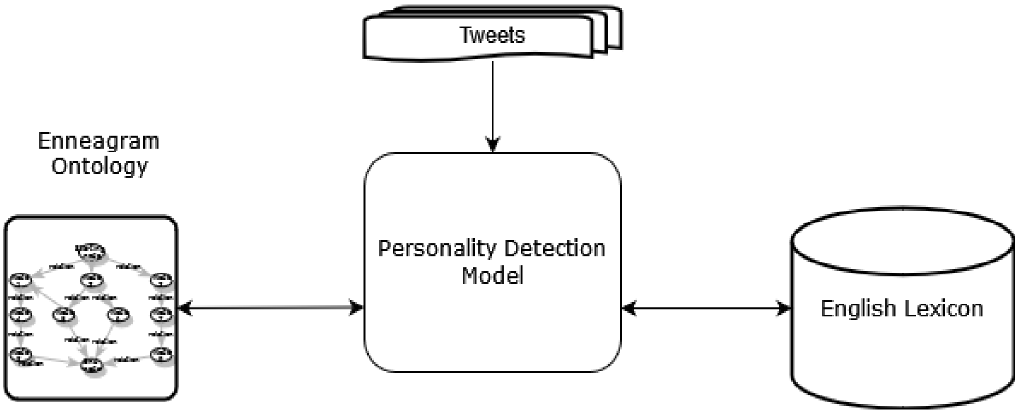

Figure 1. Personality detection model

The personality detection approach composes of several phases: text pre-processing, feature extraction, feature selection, and personality detection. The text pre-processing target is to remove unnecessary information from the text. The text pre-processing step includes text cleaning, text normalization, stemming, and lemmatization. Text cleaning involves the removal of hashtags, Twitter handlers, numbers, extra space, stop words, uniform resource locators (URLs), emails, punctuation, and special character. Text normalization composes of checking the spelling, normalizing Unicode, and lower-case folding. The last one is stemming and lemmatization words. All steps are shown in Figure 2.

Figure 2. Text preprocessing phase

In feature extraction, this stage's purpose is to obtain the qualifying features. The text is required to be converted to features. First, the preprocessed text is tokenized. Second, the tokenization is transformed into a bag of words. It contains the words as features and their relevant occurrences count.

Feature selection is vital to pick the important features. The Enneagram ontology [19], [20] and the thesaurus English lexicon [18] are both used in word-based feature selection. Ontology supplies knowledge of a domain [32]. The English lexicon enhances the vocabulary. It enriches the initial lists with equivalent words. The new word lists are formed from the initial lists which are extracted from the ontology and the equivalent words from the English lexicon. Each list is relative to one of the nine personalities. Then, the nine final word lists are preprocessed. The features are chosen from the pre-processed final words lists that are found in the bag of words. Nine bags of words are formed from the intersection between nine lists and a bag of words of the main text as shown in the next Figure 3.

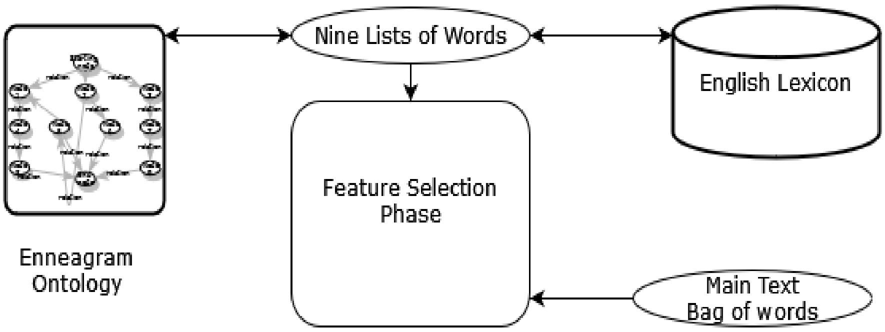

Figure 3. Word-based feature selection

Personality detection is applied to identify the nine distinct personalities. The goal of this stage is to detect the Enneagram personality. The total occurrences count for each bag of words is computed to determine the personality distribution. Every personality bag of words contains words that belong to the personality and count occurrences in the main text. The probability for a personality equals the sum of a personality bag of words occurrences count divided by the sum of word occurrences count for all personalities. Enneagram's personalities probability distribution is computed for each of the nine personalities. The highest probability distribution for a personality means that it is the detected one [21].

4. PROPOSED PERSONALITY DETECTION CASE STUDY

The proposed case study is applied to the public profile Twitter of Morgan Freeman [23]. The case study contains the biography description text and 40 tweets. The data are available here [33]. The procedure for the approach contains four steps. The steps include text pre-processing, feature extraction, feature selection, and personality detection. In the text pre-processing phase, the unnecessary text is removed. Also, the text is prepared for the next stage. The description and 40 tweets are preprocessed. The pre-processing involves removing emails, URLs, hashtags, Twitter handler, punctuations, stop words, extra spaces, special characters, and numbers. Text preprocessing also consists of lower-case folding, normalizing Unicode, checking the spelling, stemming, and lemmatization. The preprocessed text is available here. The text is ready for the next phase.

In the feature extraction phase, this stage's target is to convert preprocessed text to word-based features. The preprocessed description text and preprocessed 40 tweets are tokenized. Then, the tokenization is transformed into a bag of word-based features. The bag of words contains words in the text and the occurrences count of each of these words. There are many words and their counter occurrences like prison: 12, friend: 7, thank: 6, year: 6, new: 5, love: 4. For example, the love word is found in the preprocessed text 4 times. The details are shown in the next Figure 4.

In the feature selection phase, the objective of this phase is to elect the right candidate features. The election criteria are based on the Enneagram ontology and the English lexicon. The word list is created from the Enneagram ontology. Using the English lexicon, the list of words is updated with equivalent words. Then, the final word list is pre-processed. The features are selected from the bag of words using the final words list.

Personality detection is the last phase. The purpose of this stage is to identify a person's personality. The personality's words count occurrences are calculated for the nine personalities. Enneagram personalities' probability distributions are computed. The highest probability distribution personality is the desired one.

The largest percentage distribution is the peacemaker which is 15.58%. This result indicates that the profile's personality is the peacemaker. This result matches the determination of the Enneagram's experts [24].

'prison': 12, 'i': 9, 'friend': 7, 'join': 7, 'tonight': 7, 'story': 7, 'break': 7, 'channel': 7, 'one': 6, 'thank': 6, 'year': 6, 'alway': 5, 'new': 5, 'film': 5, 'episode': 5, 'world': 4, 'kenneth': 4, 'chamberlain': 4, 'love': 4, 'two': 3, 'make': 3, 'see': 3, 'day': 3, 'well': 3, 'way': 3, 'tomorrow': 3, 'continue': 3, 'share': 3, 'escape': 3, 'correct': 3, 'fail': 3, 'breakout': 3, 'work': 3, 'attempt': 3, 'premier': 3, 'behind': 3, 'pull': 3, 'so': 3, 'team': 3, 'novemb': 3, 'unbelief': 3, 'actor': 2, 'take': 2, 'say': 2, 'set': 2, 'like': 2, 'amaz': 2, 'rita': 2, 'weekend': 2, 'award': 2, 'want': 2, 'wish': 2, 'happi': 2, 'birthday': 2, 'busi': 2, 'partner': 2, 'compass': 2, 'great': 2, 'serv': 2, 'rememb': 2, 'man': 2, 'think': 2, 'betty': 2, 'fall': 2, 'everyone': 2, 'pleas': 2, 'inspire': 2, 'draw': 2, 'walk': 2, 'talent': 2, 'effort': 2, 'explore': 2, 'since': 2, 'look': 2, 'back': 2, 'enjoy': 2, 'larger': 2, 'project': 2, 'ill': 2, 'unfold': 2, 'inbred': 2, 'next': 2, 'disco': 2, 'war': 2, 'creat': 2, 'tuesday': 2, 'kotori': 2, 'hbo': 2, 'max': 2, 'death': 2, 'th': 2, 'thanksgiv': 2, 'inmate': 2, 'for': 2, 'clinton': 2, 'dannemora': 2, 'air': 2, 'bring': 2, 'season': 2, 'still': 2, 'artist': 2, 'digit': 2, 'mohacyucel': 2, 'portrait': 2, 'done': 2, 'whose': 1, 'voice': 1, 'recon': 1, 'better': 1, 'include': 1, 'seri': 1, 'quick': 1, 'doubt': 1, 'smile': 1, 'peopl': 1, 'wonder': 1, 'your': 1, 'guess': 1, 'who': 1, 'laugh': 1, 'old': 1, 'past': 1, 'aii': 1, 'luncheon': 1, 'miracle': 1, 'son': 1, 'miracl': 1, 'bruce': 1, 'almighty': 1, 'dear': 1, 'friendship': 1, 'humor': 1, 'cheer': 1, 'mani': 1, 'airline': 1, 'connect': 1, 'country': 1, 'reconnect': 1, 'super': 1, 'bowl': 1, 'commerce': 1, 'nelson': 1, 'mandela': 1, 'release': 1, 'madiba': 1, 'courage': 1, 'determin': 1, 'chang': 1, 'south': 1, 'africa': 1, 'last': 1, 'vega': 1, 'robert': 1, 'noth': 1, 'movi': 1, 'toget': 1, 'deserve': 1, 'vote': 1, 'call': 1, 'write': 1, 'life': 1, 'precious': 1, 'meet': 1, 'along': 1, 'heavenly': 1, 'white': 1, 'treasure': 1, 'beauty': 1, 'gloria': 1, 'foster': 1, 'milder': 1, 'would': 1, 'nature': 1, 'rather': 1, 'play': 1, 'coriolanu': 1, 'word': 1, 'kill': 1, 'sidney': 1, 'guid': 1, 'light': 1, 'send': 1, 'joanna': 1, 'family': 1, 'parti': 1, 'angel': 1, 'that': 1, 'legend': 1, 'differ': 1, 'se': 1, 'hilton': 1, 'art': 1, 'line': 1, 'paul': 1, 'knee': 1, 'dankeschon': 1, 'dominiquesfarbmanufakturink': 1, 'ig': 1, 'almost': 1, 'might': 1, 'daydream': 1, 'elvira': 1, 'success': 1, 'need': 1, 'lot': 1, 'accord': 1, 'plan': 1, 'shock': 1, 'theater': 1, 'later': 1, 'nail': 1, 'million': 1, 'dollar': 1, 'babi': 1, 'ive': 1, 'clint': 1, 'hikari': 1, 'stake': 1, 'euros': 1, 'labor': 1, 'start': 1, 'imprison': 1, 'often': 1, 'spur': 1, 'ingenue': 1, 'nitric': 1, 'tunnel': 1, 'system': 1, 'even': 1, 'possible': 1, 'entrap': 1, 'massive': 1, 'stone': 1, 'wall': 1, 'pittsburgh': 1, 'state': 1, 'penitentiary': 1, 'six': 1, 'men': 1, 'dare': 1, 'impose': 1, 'dont': 1, 'embarrass': 1, 'el': 1, 'chap': 1, 'drug': 1, 'lord': 1, 'orchestra': 1, 'legendary': 1, 'stun': 1, 'uncov': 1, 'stream': 1, 'link': 1, 'today': 1, 'ten': 1, 'entir': 1, 'honor': 1, 'office': 1, 'avail': 1, 'already': 1, 'seen': 1, 'the': 1, 'anniversary': 1, 'let': 1, 'week': 1, 'conside': 1, "don't": 1, 'lori': 1, 'event': 1, 'decadeslong': 1, 'commun': 1, 'visit': 1, 'profile': 1, 'help': 1, 'ny': 1, 'find': 1, 'pictur': 1, 'remind': 1, 'kid': 1, 'fly': 1, 'u': 1, 'salut': 1, 'veteran': 1, 'medic': 1, 'service': 1, 'simple': 1, 'believ': 1, 'watch': 1, 'history': 1, 'sneak': 1, 'peek': 1, 'tune': 1, 'inside': 1, 'infamy': 1, 'alcatraz': 1, 'good': 1, 'luck': 1, 'first': 1, 'game': 1, 'go': 1, 'photo': 1, 'magic': 1, 'bell': 1, 'isl': 1, 'fond': 1, 'memory': 1, 'cast': 1, 'crew': 1, 'sing': 1, 'televise': 1, 'jingle': 1, 'rob': 1, 'reiner': 1, 'enter': 1, 'reintroduce': 1, 'focus': 1, 'reveal': 1, 'truth': 1, 'tell': 1, 'authen': 1, 'len': 1, 'hope': 1, 'charact': 1, 'relax': 1, 'stori': 1, 'circa': 1, 'alf': 1, 'arsenio': 1, 'what': 1, 'favorit': 1, 'moment': 1, 'technology': 1, 'given': 1, 'within': 1, 'medium': 1, 'congrats': 1, 'gotham': 1, 'nomine': 1, 'commit': 1, 'role': 1, 'raise': 1, 'revel': 1, 'cant': 1, 'wait': 1, 'celebs': 1, 'im': 1, 'grandson': 1, 'alaric': 1, 'ink': 1, 'shawnoink': 1

Figure 4. Case study: bag of words

5. RESULT AND DISCUSSION

Each personality has related words that intersect with the main text's bag of words. The nine personalities' probability distributions are calculated and range from 15.58% to 6.71%. Different personalities' probability distributions are peacemaker: 15.58%, individualist: 12.12%, reformer: 11.90%, achiever: 11.69%, investigator: 11.47%, enthusiast: 11.04%, helper: 9.96%, loyalist: 9.52%, and challenger: 6.71%. Peacemaker is the largest percentage produced. Many features are selected with their count for the peacemaker like 'take', 'recon', 'set', 'happi', 'pleas', 'creat', 'hope'. Each selected feature has a count for each occurrence like 'take': 2, 'recon': 1,'set': 2, 'happi': 2, 'pleas': 2, 'hope': 1.

Peacemaker selected features with their count are the following: 'take': 2, 'recon': 1, 'set': 2, 'happi': 2, 'pleas': 2, 'disco': 2, 'make': 3, 'plan': 1, 'start': 1, 'creat': 2, 'relax': 1, 'seen': 1, 'last': 1, 'hope': 1, 'help': 1, 'cheer': 1, 'good': 1, 'like': 2, 'talent': 2, 'back': 2, 'inspire': 2, 'raise': 1, 'continue': 3, 'busi': 2, 'commit': 1, 'friendship': 1, 'toget': 1, 'tune': 1, 'determin': 1, 'fail': 3, 'fall': 2, 'differ': 1, 'break': 7, 'accord': 1, 'love': 4, 'meet': 1, 'connect': 1, 'join': 7, 'link': 1 . The sum of the selected peacemaker features count is 72. The sum of all personalities selected features count is 462. Peacemaker's probability distribution equals 72 divided by 462 so the computed result is 15.58 %. The illustration for each personality, selected words-based features with count and the probability distribution percentage are shown in the Table 2.

The highest probability distribution is the peacemaker personality with 15.58%. The result states that the profile personality is a peacemaker. The outcome matches the Enneagram experts' analysis [24]. This result drives us to the conclusion that the approach can determine the Enneagram's personality. The integration of the Enneagram ontology, English lexicon, and statistical technique are qualified for detecting personality. The author's profile text gives the direction of his/her personality.

Current research on personality detection utilizes the MBTI, the big five model, and the three-factor model. Enneagram is more difficult and deeply perceives the personality. Other personality models assess specific traits without providing insight into behavior's motivation. The Enneagram not only outlines all personality characteristics but also provides every personality's deeply rooted patterns of thoughts, feelings, and behaviors [34]. This is the first Enneagram personality detection approach. This approach encourages more research in Enneagram personality detection. In the future, applying Enneagram personality identification will lead to many benefits. This minimizes time and effort for knowing their personality.

Table 2. Morgan Freeman's case study result

Personality type	Selected features	Probability percentage
Peacemaker	'take': 2, 'recon': 1, 'set': 2, 'happi': 2, 'pleas': 2, 'disco': 2, 'make': 3, 'plan': 1, 'start': 1, 'creat': 2, 'relax': 1, 'seen': 1, 'last': 1, 'hope': 1, 'help': 1, 'cheer': 1, 'good': 1, 'like': 2, 'talent': 2, 'back': 2, 'inspire': 2, 'raise': 1,'continue': 3, 'busi': 2, 'commit': 1, 'friendship': 1, 'toget': 1, 'tune': 1, 'determin': 1, 'fail': 3, 'fall': 2, 'differ': 1, 'break': 7, 'accord': 1, 'love': 4, 'meet': 1, 'connect': 1, 'join': 7, 'link': 1	15.58%
Individualist	'beauty': 1, 'commun': 1, 'reveal': 1, 'say': 2, 'tell': 1, 'voice': 1, 'compass': 2, 'draw': 2, 'like': 2, 'pictur': 1, 'portrait': 2, 'state': 1, 'man': 2, 'parti': 1, 'air': 2, 'charact': 1, "humor': 1, 'wish': 2, 'recon': 1, 'release': 1,'disco': 2, 'make': 3, 'plan': 1, ' 'start': 1, 'creat': 2, 'spur': 1, 'inspire': 2, 'continue': 3, 'play': 1, 'theater': 1,'authen': 1, 'honor': 1, 'since': 2, 'embarrass': 1, 'believ': 1, 'think': 2, 'nature': 1, 'moment': 1, 'differ': 1	12.12%
Reformer	'differ': 1, 'join': 7, 'link': 1, 'great': 2, 'wonder': 1,'talent': 2, 'help': 1, 'good': 1, 'honor': 1, 'line': 1, 'plan': 1,'system': 1, 'tell': 1, 'creat': 2, 'simple': 1, 'truth': 1, 'determin': 1, 'project': 2, 'wish': 2, 'better': 1, 'correct': 3,'raise': 1, 'recon': 1, 'happi': 2, 'busi': 2, 'ill': 2, 'fall': 2, 'impose': 1, 'authen': 1, 'since': 2, 'war': 2, 'success':1, 'continue': 3, 'cheer': 1	11.90%
Achiever	'honor': 1, 'treasure': 1, 'think': 2, 'determin': 1, 'recon': 1, 'disco': 2, 'see': 3, 'inspire': 2, 'authen': 1, 'spur': 1,'draw': 2, 'like': 2, 'pictur': 1, 'portrait': 2, 'state': 1, 'hope': 1, 'love': 4, 'wish': 2, 'bring': 2, 'kill': 1, 'send': 1,'play': 1, 'beauty': 1, 'magic': 1, 'pleas': 2, 'effort': 2, 'for': 2, 'make': 3, 'fly': 1, 'guid': 1, 'start': 1, 'service':1, 'seen': 1, 'wait': 1, 'courage': 1, 'favorit': 1	11.69%
Investigator	'determin': 1, 'explore': 2, 'recon': 1, 'reveal': 1, 'see': 3, 'disco': 2, 'pictur': 1, 'war': 2, 'watch': 1, 'wonder': 1, 'plan': 1, 'start': 1, 'great': 2, 'find': 1, 'nail': 1, 'good': 1, 'talent': 2, 'play': 1, 'back': 2, 'conside': 1, 'break': 7,'tell': 1, 'guess': 1, 'include': 1, 'project': 2, 'call': 1, 'think': 2, 'ill': 2, 'service': 1, 'serv': 2, 'help': 1, 'need': 1, 'work': 3	11.47%
Enthusiast	'believ': 1, 'don't': 1, 'work': 3, 'busi': 2, 'game': 1, 'set': 2, 'take': 2, 'play': 1, 'system': 1, 'simple': 1, 'talent':2, 'enjoy': 2, 'recon': 1, 'honor': 1, 'like': 2, 'love': 4, 'treasure': 1, 'meet': 1, 'focus': 1, 'cheer': 1, 'humor':1, 'wonder': 1, 'happi': 2, 'pleas': 2, 'chang': 1, 'charact': 1, 'nature': 1, 'rob': 1, 'ill': 2, 'shock': 1, 'amaz': 2,'event': 1, 'avail': 1, 'release': 1, 'success': 1, 'help': 1	11.04%
Helper	'enjoy': 2, 'recon': 1, 'honor': 1, 'like': 2, 'love': 4, 'treasure': 1,'dear': 1, 'relax': 1, 'meet': 1, 'commun': 1,'reveal': 1, 'say': 2, 'tell': 1, 'voice': 1, 'service': 1, 'back': 2,'serv': 2, 'help': 1, 'commit': 1, 'wish': 2, 'want': 2,'need': 1, 'effort': 2, 'watch': 1, 'determin': 1, 'compass': 2, 'good': 1, 'cheer': 1, 'pleas': 2, 'since': 2, 'fond': 1, 'friendship': 1	9.96%
Loyalist	'send': 1, 'commit': 1, 'help': 1, 'join': 7, 'meet': 1, 'fall': 2, 'busi': 2, 'determin': 1, 'for': 2, 'quick': 1, 'doubt':1, 'war': 2, 'believ': 1, 'conside': 1, 'think': 2, 'wonder': 1, 'correct': 3, 'hope': 1, 'cheer': 1, 'inspire': 2, 'super':1, 'last': 1, 'take': 2, 'back': 2, 'raise': 1, 'continue': 3	9.52%
Challenger	'say': 2, 'meet': 1, 'accord': 1, 'determin': 1, 'find': 1, 'talent': 2, 'for': 2, 'system': 1, 'success': 1, 'good': 1,'courage': 1, 'great': 2, 'charact': 1, 'wish': 2, 'impose': 1, 'better': 1, 'correct': 3, 'help': 1, 'raise': 1, 'recon': 1, 'spur': 1, 'inspire': 2, 'humor': 1	6.71%

Automated Enneagram personality detection has an impact in different areas like recommendation systems, psychiatrists, psychologists, education. In recommendation systems, this identification recognizes the likes and dislikes of the person's profile which enhances the commercial advertising. Better recommendation systems that can act in human-like performance. Quick recognition helps the psychiatrist and psychologist to give the desired assistance to patients in less time. It can also be applied in education to provide support for students to raise their achievements. Enneagram can enhance postgraduate education, professional connections, and boost motivation during tough patient care situations [35].

The result cannot demonstrate the level of accuracy of the approach. The case study investigates a single profile that does not give the big picture. Other personalities are required to explore. More case studies need to be discussed. Different profiles with different personalities are also needed. Various profiles will be analyzed to determine the level of accuracy.

6. CONCLUSION

An Enneagram personality identification case study is proposed. The public Twitter profile text is analyzed which contains a biography description and 40 tweets. The steps of the approach are multiple: text preprocessing, feature extraction, feature selection, and personality detection. The main goal of text preprocessing is to clean the text from unimportant and additional text. Feature extraction obtains words from text as word-based features. The purpose of feature selection is to elect the right features with the combination of the Enneagram ontology and the English lexicon. The personality detection phase uses probability distribution across personalities. The largest value of the probability distributions is the desired personality.

According to the proposed case study, text pre-processing contains normalization, cleaning, lemmatization, and stemming. Feature extraction includes tokenization and bag of words representation. Feature selection is based on the list of words and bag of words. Personalities lists are formed from Enneagram ontology and English Lexicon. Personality-related features are chosen from the personality-related list of words and found in the bag of text. Personality detection is applied based on calculating the personalities' probability distributions. The probability distributions for each personality from the highest to the lowest are peacemaker: 15.58%, individualist: 12.12%, reformer: 11.90%, achiever: 11.69%, investigator: 11.47%, enthusiast: 11.04%, helper: 9.96%, loyalist: 9.52%, and challenger: 6.71%. The highest probability distribution is 15.58%. This percentage belongs to the peacemaker personality. This result inferred that the profile personality is the peacemaker personality. This result matches the Enneagram experts' conclusion.

This is the first Enneagram personality detection approach. Past research focused on other personality models. The Enneagram provides a deeper understanding because it illustrates the person's desires, motivations, and fears. The case study result is identical to the Enneagram's expert analysis. The output indicates a positive direction. This system can help recommendation systems, education, psychiatrists, psychologists, and physicians instead of applying the test. This paper is a gate and encourages more investigation in this domain.

In the future, the approach will be applied to more case studies and results. Another future enhancement, a system will be developed that can infer the person's personality in unhealthy and healthy conditions. Unhealthy conditions may lead to suicide attempts or/and personality disorders. Different approaches will be implemented to compare the outcomes.

REFERENCES

[1] M. Farhadloo and E. Rolland, "Fundamentals of sentiment analysis and its applications," in *Studies in Computational Intelligence*, vol. 639, Springer International Publishing, 2016, pp. 1–24.
[2] M. Kosinski, Y. Bachrach, P. Kohli, D. Stillwell, and T. Graepel, "Manifestations of user personality in website choice and behaviour on online social networks," *Machine Learning*, vol. 95, no. 3, pp. 357–380, 2014, doi: 10.1007/s10994-013-5415-y.
[3] O. M. Şulea and D. Dichiu, "Automatic profiling of Twitter users based on their tweets," *CEUR Workshop Proceedings*, vol. 1391, 2015.
[4] G. Farnadi *et al.*, "Computational personality recognition in social media," *User Modeling and User-Adapted Interaction*, vol. 26, no. 2–3, pp. 109–142, Jun. 2016, doi: 10.1007/s11257-016-9171-0.
[5] W. Arthur, D. J. Woehr, and W. G. Graziano, "Personality testing in employment settings: problems and issues in the application of typical selection practices," *Personnel Review*, vol. 30, no. 6, pp. 657–676, Dec. 2001, doi: 10.1108/EUM0000000005978.
[6] A. Vinciarelli and G. Mohammadi, "A survey of personality computing," *IEEE Transactions on Affective Computing*, vol. 5, no. 3, pp. 273–291, Jul. 2014, doi: 10.1109/TAFFC.2014.2330816.
[7] D. Azucar, D. Marengo, and M. Settanni, "Predicting the big 5 personality traits from digital footprints on social media: a meta-analysis," *Personality and Individual Differences*, vol. 124, pp. 150–159, Apr. 2018, doi: 10.1016/j.paid.2017.12.018.
[8] S. J. Guy, S. Kim, M. C. Lin, and D. Manocha, "Simulating heterogeneous crowd behaviors using personality trait theory," in *Proceedings - SCA 2011: ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, Aug. 2011, pp. 43–52, doi: 10.1145/2019406.2019413.
[9] D. R. Riso and R. Hudson, *The Wisdom of the Enneagram: The Complete Guide to Psychological and Spiritual Growth for the Nine Personality Types*. Bantam, 1999.
[10] A. M. Bland, "The Enneagram: a review of the empirical and transformational literature," *Journal of Humanistic Counseling, Education and Development*, vol. 49, no. 1, pp. 16–31, Mar. 2010, doi: 10.1002/j.2161-1939.2010.tb00084.x.
[11] A. Demir, O. Rakhmanov, K. Tastan, S. Dane, and Z. Akturk, "Development and validation of the Nile personality assessment tool based on Enneagram," *Journal of Research in Medical and Dental Science*, 2020.
[12] R. Baron and E. Wagele, *The Enneagram Made Easy Discover the 9 Types of People*. HarperOne, 2009.
[13] M. Matise, "The Enneagram: an enhancement to family therapy," *Contemporary Family Therapy*, vol. 41, no. 1, pp. 68–78, Jun. 2019, doi: 10.1007/s10591-018-9471-0.
[14] M. Alexander and B. Schnipke, "The Enneagram: a primer for psychiatry residents," *American Journal of Psychiatry Residents' Journal*, vol. 15, no. 3, pp. 2–5, Mar. 2020, doi: 10.1176/appi.ajp-rj.2020.150301.
[15] L. Huffman, E. M. Lefdahl-Davis, and A. Alayan, "The Enneagram and the college student: empirical insight, legitimacy, and practice," *Christian Higher Education*, vol. 21, no. 3, pp. 214–232, Jun. 2022, doi: 10.1080/15363759.2021.1929566.
[16] V. Kaushal and M. Patwardhan, "Emerging trends in personality identification using online social networks-a literature survey," *ACM Transactions on Knowledge Discovery from Data*, vol. 12, no. 2, pp. 1–30, Jan. 2018, doi: 10.1145/3070645.
[17] E. Cambria, S. Poria, A. Gelbukh, and M. Thelwall, "Sentiment analysis is a big suitcase," *IEEE Intelligent Systems*, vol. 32, no. 6, pp. 74–80, Nov. 2017, doi: 10.1109/MIS.2017.4531228.
[18] "Thesaurus English lexicon," Thesaurus.com. [Online]. Available: https://www.thesaurus.com/ (accessed Jan. 30, 2023).
[19] E. Abdelhamid, S. Ismail, and M. Aref, "Enneaontology: a proposed enneagram ontology," in *Smart Innovation, Systems and Technologies*, vol. 302, Springer Nature Singapore, 2022, pp. 559–567.
[20] E. A. Abdelhamid, S. Ismail, and M. Aref, "Enneaontology: toward an Enneagram personality detection," in *Lecture Notes in Networks and Systems*, vol. 612, Springer Nature Singapore, 2023, pp. 1–9.
[21] E. Abdelhamid, S. Ismail, and M. Aref, "Architecture for personality detection using Enneagram knowledge: case study," *International Journal of Intelligent Computing and Information Sciences*, pp. 1–11, 2023, doi: 10.21608/ijicis.2023.137912.1181.

[22] A. Java, X. Song, T. Finin, and B. Tseng, "Why we twitter: understanding microblogging usage and communities," in *Joint Ninth WebKDD and First SNA-KDD 2007 Workshop on Web Mining and Social Network Analysis*, Aug. 2007, pp. 56–65, doi: 10.1145/1348549.1348556.

[23] "Morgan freeman's twitter," *Twitter.com* [Online]. Available: https://twitter.com/morgan freeman/ (accessed May 25, 2022).

[24] "The peacemaker Enneagram type nine," *The Enneagram Institute*. Accessed: Jan. 30, 2023. [Online]. Available: https://www.enneagraminstitute.com/type-9.

[25] B. Agarwal, "Personality detection from text: a review," *International Journal of Computer System*, vol. 1, no. 1, pp. 1–4, 2014.

[26] V. Ong, A. D. S. Rahmanto, Williem, and D. Suhartono, "Exploring personality prediction from text on social media: a literature review," *Internetworking Indonesia Journal*, vol. 9, no. 1, pp. 65–70, 2017.

[27] M. Gjurković, I. Vukojević, and J. Šnajder, "SIMPA: statement-to-item matching personality assessment from text," *Future Generation Computer Systems*, vol. 130, pp. 114–127, May 2022, doi: 10.1016/j.future.2021.12.014.

[28] L. Zhou, Z. Zhang, L. Zhao, and P. Yang, "Attention-based BiLSTM models for personality recognition from user-generated content," *Information Sciences*, vol. 596, pp. 460–471, Jun. 2022, doi: 10.1016/j.ins.2022.03.038.

[29] A. Alamsyah, M. F. Rachman, C. S. Hudaya, R. P. Putra, A. I. Rifkyano, and F. Nurwianti, "A progress on the personality measurement model using ontology based on social media text," in *Proceedings of 2019 International Conference on Information Management and Technology, ICIMTech 2019*, Aug. 2019, pp. 581–586, doi: 10.1109/ICIMTech.2019.8843817.

[30] A. Alamsyah, S. Widiyanesti, R. D. Putra, and P. K. Sari, "Personality measurement design for ontology based platform using social media text," *Advances in Science, Technology and Engineering Systems*, vol. 5, no. 3, pp. 100–107, 2020, doi: 10.25046/aj050313.

[31] A. Alamsyah, N. Dudija, and S. Widiyanesti, "New approach of measuring human personality traits using ontology-based model from social media data," *Information (Switzerland)*, vol. 12, no. 10, Oct. 2021, doi: 10.3390/info12100413.

[32] N. F. Noy and D. L. McGuinness, "Ontology development 101: a guide to creating your first ontology," *Sustainability (Switzerland)*, vol. 9, no. 12, pp. 1–25, 2017.

[33] EAAcs, "Morgan freeman's data," GitHub, https://github.com/EAAcs/Morgan-Freeman-Twitter-Dataset/blob/main/Morgan Freeman Twitter Text.xlsx (accessed May 25, 2022).

[34] H. R. Roh *et al.*, "Understanding medical students' empathy based on Enneagram personality types," *Korean Journal of Medical Education*, vol. 31, no. 1, pp. 73–82, Mar. 2019, doi: 10.3946/kjme.2019.120.

[35] T. M. Blose, A. C. Yeates, M. Som, K. A. Murray, M. Vassar, and J. Stroup, "The Enneagram and its application in medical education," *Baylor University Medical Center Proceedings*, vol. 36, no. 1, pp. 54–58, 2023, doi: 10.1080/08998280.2022.2132591.

BIOGRAPHIES OF AUTHORS

Esraa Abdelhamid is a Ph.D. student in the computer science department at the faculty of computer and information science, Ain Shams University, Cairo, Egypt. She received her master's degree in computer science from the faculty of computer and information science, Ain Shams University in 2016. Her research interests include text mining, natural language processing, data mining, knowledge representation, and multiagent systems. She can be contacted at email: esraa_cs@hotmail.com.

Sally S. Ismail holds a Ph.D. in Computer Science field from the University of Ain Shams, subspecialty in natural language processing. She is currently a lecturer in the Computer Science Department, Faculty of Computer and Information Sciences, and formerly the program coordinator in the software engineering program. Her research interests include information retrieval, text summarization, and sentiment analysis. She can be contacted via email: Sallysaad@cis.asu.edu.eg.

Mostafa Aref is a professor of Computer Science at Ain Shams University, Cairo, Egypt. He holds a Ph.D. in engineering science in system Theory and engineering, June 1988, University of Toledo, Toledo, Ohio. He obtained his M.Sc. in computer science, October 1983, University of Saskatchewan, Saskatoon, Sask. Canada. He received the B.Sc. of electrical engineering-computer and automatic control section, in June 1979, electrical engineering department, Ain Shams University, Cairo, Egypt. His research areas are natural language processing, knowledge representation, and ontology. He can be contacted at email: mostafa.m.aref@gmail.com.