International Journal of Computer Integrated Manufacturing

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tcim20

Digital twins in human understanding: a deep learning-based method to recognize personality traits

Jianshan Sun, Zhiqiang Tian, Yelin Fu, Jie Geng & Chunli Liu

To cite this article: Jianshan Sun, Zhiqiang Tian, Yelin Fu, Jie Geng & Chunli Liu (2021) Digital twins in human understanding: a deep learning-based method to recognize personality traits, International Journal of Computer Integrated Manufacturing, 34:7-8, 860-873, DOI: 10.1080/0951192X.2020.1757155

To link to this article: https://doi.org/10.1080/0951192X.2020.1757155

INTERNATIONAL JOURNAL OF COMPUTER INTEGRATED MANUFACTURING
2021, VOL. 34, NOS. 7–8, 860–873
https://doi.org/10.1080/0951192X.2020.1757155

ARTICLE

Digital twins in human understanding: a deep learning-based method to recognize personality traits

Jianshan Sun[a,b], Zhiqiang Tian[a], Yelin Fu[c,d], Jie Geng[a] and Chunli Liu[a]

[a]School of Management, Hefei University of Technology, Hefei, China; [b]Key Laboratory of Process Optimization and Intelligent Decision Making, Ministry of Education, Hefei, Anhui, China; [c]School of Intelligent Systems Science and Engineering, Jinan University (Zhuhai Campus), Zhuhai, China; [d]Department of Industry and Manufacturing Systems Engineering, The University of Hong Kong, Hong Kong, China

ABSTRACT

Digital twin models are computerized clones of physical assets or systems and have attracted much attention from academia and industries. Digital twin applications focus on smart manufacturing systems. Meanwhile, manufacturing products are driven increasingly by the needs of customers. Industrial production modes have evolved from mass production to personalized production. Understanding customers and meeting their personalized needs have become important issues in smart manufacturing. Social networks provide platforms for online customers to engage in different behaviors. In addition, personality recognition is a crucial issue for understanding people. In this study, a new technique is proposed to formalize personality as digital twin models by observing users' posting content and liking behavior. A multitask learning deep neural network model is used to predict users' personality through two types of data representation. Experimental results show that combining the two types of data can improve personality prediction accuracy.

ARTICLE HISTORY
Received 22 April 2019
Accepted 23 March 2020

KEYWORDS
Digital twin; personality recognition; deep neural network; multitasking learning; big five personality

1. Introduction

Developed originally to improve manufacturing processes, a digital twin is a virtual representation of a physical asset or system throughout its lifecycle (Tao et al. 2018). The concept of digital twins is applied broadly to numerous fields beyond manufacturing, such as health and wellness, security and safety, transport and energy, and mobility and communications (El Saddik 2018; Bruynseels, Santoni de Sio, and Hoven 2018; Banerjee et al. 2017). Digital twins enable the seamless transmission of data between physical and virtual systems and facilitate the means to monitor, understand, and optimize the functions of physical living and nonliving entities (El Saddik 2018). Recently, digital twin technology has been utilized to improve human health and wellbeing (Bruynseels, Santoni de Sio, and Hoven 2018). It can show what is happening inside the bodies of real twins and predict illnesses by analyzing the real twins' data. Beyond health and wellbeing, digital twin technology can be used to understand human characteristics and provide highly accurate descriptions of reals twin at different privacy levels. The current work focuses on personality (which is a specific human characteristic) and creates a digital twin of a personality in online social networks.

Presently, various personality trait theories are applied to different scenarios, such as business management, clinical trials, marketing, and personalized recommendations. For example, Blignaut and Annelie (2008) focused on a computer course and found that students' personality traits could affect their classroom performance. Judge et al. (1999) devoted themselves to investigating the intrinsic link between the Big Five personality traits and career success. In business areas, to attract additional customers, Chen, Tsai, and Chen (2016) developed a system for predicting individual personality types and designed adaptive marketing strategies based on recognized user personalities. Meanwhile, according to Knapp and Daly (2002), personality determines how people behave and interact with others.

Online social networks provide platforms for people to behave and communicate with others. These platforms can be considered as virtual reflections of human behaviors where behavior data are recorded. On these platforms, people can not only express their preferences freely but also participate in interpersonal

CONTACT Yelin Fu msylfu@gmail.com Pokfulam, Hong Kong, China

activities. In such processes, people can demonstrate their language, behavior, psychology, and other aspects. Therefore, analyzing personality characteristics with social network data is feasible.

In previous research, user-generated textual content is employed to predict individual personality traits. Mairesse et al. (2007) pointed out that individual traits can be recognized based on language cues from textual content. Moreover, users' interaction behavior data can be used for personality prediction. Li et al. (2011) indicated that a strong relationship exists between users' online behaviors and personality traits as well as other psychological characteristics. Furthermore, digital footprints have been proven effective in predicting individuals' personality (Wu, Kosinski, and Stillwell 2015). In recent years, deep learning technologies have been applied to numerous fields, especially document modeling and natural-language processing. Therefore, researchers have employed neural networks in personality prediction tasks. Zhang et al. (2017b) used a convolutional neural network to learn individual facial expression features then leveraged the features to build a personality prediction model.

Although neural network technology has been applied to personality prediction models, no work has combined individuals' text and behavioral data with neural network technology. The present work will employ users' text content and liking behavior digital footprints through a neural network. Figure 1 presents an example of the user data employed in this work. A word-embedding model is used to obtain users' document vector. On this basis, users' likes vector modeled by singular value decomposition (SVD) as input is fed into the deep neural network (DNN). The experiments show that the use of like behavior data can improve user personality prediction accuracy. Furthermore, through further analysis, we find that the higher the likes record, the higher the accuracy of the model.

The rest of the paper is organized as follows. Section 2 presents previous research on personality prediction, and Section 3 introduces the proposed method, which is a multitask DNN model for Big Five personality prediction. The specific details of the experiments and result analysis are described

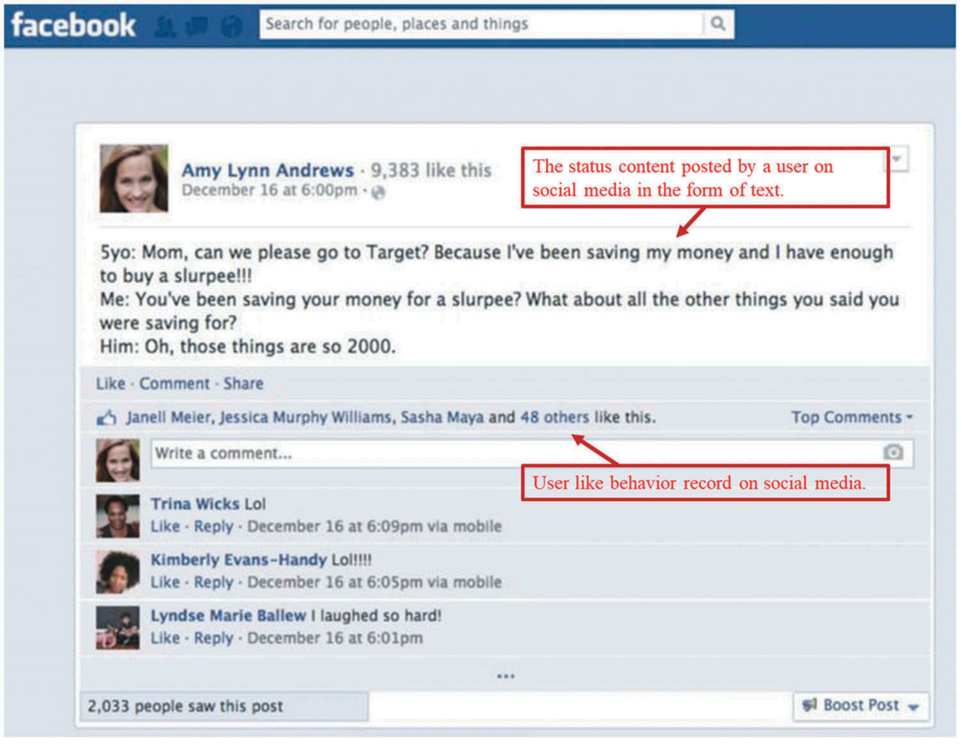

Figure 1. Example of user data.

in the Section 4. Finally, the study is summarized, and future research directions are identified in Section 5.

2. Related research

2.1. Digital twin in manufacturing systems

Increasing amounts of data are generated in different stages of the manufacturing industry, including product design, raw material procurement, warehousing, production, sales, and services (Li et al. 2015). Data include not only machine data generated by manufacturing equipment and sensors in a manufacturing plant but also human data generated through production planning, quality assessment, and consumer feedback. Managing data to improve the performance of the manufacturing industry is an important issue for enterprises and researchers. Digital twins integrate physical and virtual products, connect data in physical and virtual spaces, and enable the fusion of different space data (Tao et al. 2018a). Concurrently, digital twin technology relies on expert knowledge and the collected data during a product's life cycle to improve and update products (Tao et al. 2018b).

In the product design stage, digital twins are employed often to enhance the efficiency and responsiveness of the design process. Tao et al. (2019) proposed a DT-driven design framework to facilitate product design decisions. Moreover, the framework can be applied to different design stages, including product planning, conceptual design, and detailed design. Schleich et al. (2017) proposed a new product DT model to realize geometrical variation management, whereas Söderberg et al. (2017) explored DT-based geometry assurance for individualized product production. Compared with a single product, the product line design process is monitored and realized through a DT-based approach, which is evaluated in the case of glass product lines (Zhang et al. 2017a).

In the product operation stage, digital twins are used to enhance the reliability and predictability of the production process. In the context of smart manufacturing, a new DT-based architecture design approach is proposed for the modularized design of cyber physical production systems (Stark, Kind, and Neumeyer 2017). Given that connected micro smart factories in factory-as-a-service systems are inefficient in terms of cost and production, Park et al. (2019) designed and implemented a digital twin application for a connected micro smart factory and realized the efficient operation of a manufacturing system. In a manufacturing system, reconfigurations are often manually operated to meet new production requirements. Zhang et al. (2019) proposed a five-dimensional fusion model for a virtual entity digital twin for robotics-based smart manufacturing systems to support automatic reconfiguration.

In the prognostic and health management (PHM) stage, digital twins are utilized to enhance model accuracy and robustness and applied extensively to equipment accident detection (Smarslok, Culler, and Mahadevan 2012), equipment life prediction (Tuegel et al. 2011), and effective analysis (Bielefeldt, Hochhalter, and Hartl 2015). Given that PHM is crucial in the lifecycle monitoring of a product, Tao et al. (2018a) first constructed a general DT then proposed a new DT-driven PHM method for complex equipment. In addition, the authors employed a case study on a wind turbine to illustrate the effectiveness of the proposed method. To enable predictive maintenance for manufacturing resources, Aivaliotis, Georgoulias, and Chryssolouris (2019) applied physics-based simulation models and the digital twin concept to calculate the remaining useful life of machinery and equipment. Furthermore, Hochhalter et al. (2014) combined digital twin technology with sensory materials to improve the prediction accuracy of the repair and replacement process. The effectiveness of DT is also demonstrated by a case study on nonstandard specimens.

Existing digital twins are employed mainly in manufacturing and service phases. However, research on how to use the user feedback information of virtual spaces to improve product design remains limited. The current work focuses on data generated by consumers of products as well as the products' corresponding target markets to provide timely feedback to the manufacturing industry.

2.2. Personality measurement

Psychology studies show that personality traits can explain individuals' behaviors and preferences to a large extent (Ozer and Benet-Martinez 2006). Funder (2001) found that personality can reflect people's consistent behavior patterns, such as behavior styles, ways of thinking, and interpersonal interactions. Similarly, Wu, Kosinski, and Stillwell (2015) indicated that personality is the key driver of people's interactions, behaviors, and emotions.

2.2.1. Myers–briggs type indicator (MBTI)

The MBTI was proposed by psychologists Myers et al. (2003). The theory holds that an individual's personality can be analyzed from four dimensions, that is, source of driving force, way of accepting information, way of making decisions, and attitude toward uncertainty. Each dimension can be divided into two opposite poles, and two dimensions can be combined to form 16 personality types. This theory is used mainly to help individuals understand a subject's work style, career adaptability, and so on.

2.2.2. Dominance, influence, robustness, and caution (DISC)

The DISC model is based on four basic behavioral style factors, namely, dominance, influence, robustness, and caution, proposed by American psychologist Marston (2013). Given that these behavioral style factors are identifiable and observable, the DISC model is suitable for various commercial fields. For example, numerous companies adopt this theory for promotional activities. Specifically, to increase the likelihood of trading, salesmen are trained to judge customers' DISC personality and correspondingly adjust their verbal marketing strategy (Chen, Tsai, and Chen 2016).

2.2.3. Big five personality

McCrae and Jr (1997) believed that individuals' personality can be measured through five factors, namely, openness (O), conscientiousness (C), extroversion (E), pleasantness (A), and neuroticism (N). The model summarizes every aspect of personality and quantifies personality by scoring the five factors. Currently, the Big Five model is considered the most popular and widely accepted personality structure. Therefore, the current work chooses the Big Five model for personality prediction.

2.3. Personality prediction

From the nature of the task, personality trait prediction can be divided into two categories, namely, classification and regression. The classification task classifies each personality trait dimension into different levels (i.e. high, medium, and low). The regression task predicts the score of each trait dimension and understands differences among individuals. Scholars have used the classification method to study personality traits. Mønsted, Mollgaard, and Mathiesen (2018) divided the scores of participants' personality traits into three levels, using 0, 1, and 2 to represent the different levels of the score. In addition, scholars have employed support vector machines (SVMs) with a radial basis function kernel (Hearst et al. 1998) to predict classification labels for individual personalities. Hassanein et al. (2018) constructed two vectors to represent each dimension of the Big Five personality traits. Moreover, the authors used a classification method for prediction and incorporated semantic similarities for consideration.

Furthermore, numerous studies have utilized regression models to obtain accurate personality trait scores. Nave et al. (2018) built a linear regression (LR) with LASSO penalty based on individuals' music preferences. Similarly, Baik et al. (2016) and Samani et al. (2018) established an LR model based on users' interaction logs, such as relationships or liking behavior. Park et al. (2015) established a ridge regression prediction model based on the word frequency statistics of a text. By contrast, Liu, Wang, and Jiang (2016) presented a novel model, namely, the personality-trait LDA. The authors' model considers the various topics extracted from user-generated content and converts thousands of N-gram features into potential topics.

With the development of online social networks, data sources have become massive, diverse, and complex. Faced with such scenarios, using data effectively has become crucial and challenging. To improve personality prediction, the AAAI conference specially organized a discussion on how to choose data features and models to adopt. From the perspective of feature types, research on personality traits and social networks is divided mainly into two types, that is, the relationship between personality traits and individual behaviors and the relationship between personality traits and user-generated content. User behavior in social networks is complex and diverse, including establishing friendships, publishing original post information, commenting on other people's post content, using other applications, and so on. User behavior in social networks is divided into five categories

(Ross et al. 2009), namely, posting, forwarding, commenting, liking, and attracting attention. Therefore, numerous scholars have explored user behavior. Amichai-Hamburger and Vinitzky (2010) emphasized that certain traits are closely related to individuals' usage patterns on Facebook. Chen, Tsai, and Chen (2016) built a prediction model with various quantitative statistical features of interaction behavior. Moreover, individuals' personality traits, emotions, or moods can be inferred from user-generated content (Back et al. 2010; Farnadi et al. 2013; Golbeck, Robles, and Turner 2011).

Literature exists on the mining of text content to predict personality. The Linguistic Inquiry and Word Count (LIWC) tool is a well-known text analysis software (Pennebaker and King 1999). Through LIWC, researchers can extract several statistical features that are related to standard counts, psychological processes, personal concerns, or linguistic dimensions. Meanwhile, the features of Structured Programming for Linguistic Cue Extraction (SPLICE) contain not only hits on whether self-assessment is positive or negative but also complex and reliable scores (Moffitt et al. 2010). NRC is an emotion lexicon that includes more than 14,000 English words depicting eight emotions, namely, anger, fear, anticipation, trust, surprise, sadness, joy, and disgust, and two sentiments, that is, positive and negative (Mohammad, Zhu, and Martin 2014). Moreover, MRC is a psycholinguistic database that depicts the psychological and distributional information of words. Specifically, more than 150,000 entries exist with information on 26 properties, such as the phoneme counts of words. These features are used frequently in personality prediction tasks. Farnadi et al. (2016) combined the feature sets of LIWC, MR, and SPLICE with individuals' interaction behaviors to predict personality impression scores. Tadesse et al. (2018) predicted individuals' personality traits with word embedding and the feature sets of SPLICE and LIWC. To obtain improved prediction results, researchers have explored various methods, such as LR, SVM, Random Forest, LDA, and so on. For example, Xue et al. (2018) designed a two-level hierarchical DNN model by introducing the attention mechanism (AttRCNN–CNNs) and incorporated semantic and statistical language features into a traditional regression algorithm to predict personality scores. Gavrilescu and Vizireanu (2017) likewise proposed a novel noninvasive neural network-based architecture to predict personality by analyzing handwriting.

Although considerable research has been conducted to predict users' personality traits from different aspects, the use of multiple behavioral data to predict users' Big Five personality is rare. The current study focuses on the fusion of textual data and liking behavior record and employs multitask learning to predict users' Big Five personality simultaneously. Multitask learning is widely used in numerous fields, such as image processing (Zhang et al. 2016) and natural-language processing (Collobert and Weston 2008), to use the same model to accomplish different tasks. Given that the Big Five personality traits are considered as a whole when describing people's characteristics, multitask learning is appropriate for personality prediction.

3. Methodology

In the product design phase of manufacturing, designers must consider various data and information, such as devices, product sales, and consumer feedback data. These data are huge and scattered in different places. Digital twin technology facilitates the integration of data to guide designers in updating their product designs quickly. The current study focuses on data generated by consumers on social media. Through data mining, changes in a target market can be fed back immediately to a company to guide product development. We apply the communication framework between a digital twin and a real twin developed by El Saddik (2018) to the context of human understanding and personality recognition (Figure 2). The digital twins in this framework have several characteristics. (1) Sensing and actuating: Real twins can be equipped with sensors to replicate their senses, namely, sight, hearing, taste, smell, and touch, using appropriate actuators depending on application needs. In online social networks, the posting content and liking behavior of users can be recorded to sense and understand users' personality. (2) Intelligence: Digital twins should be equipped with a controller embedded with deep learning techniques to make fast and intelligent decisions on behalf of their real twin. In this study, we leverage deep learning techniques to combine users' posting content and liking behavior for personality prediction. (3) Representation: Digital twins can be represented virtually as a social humanoid robot or a software

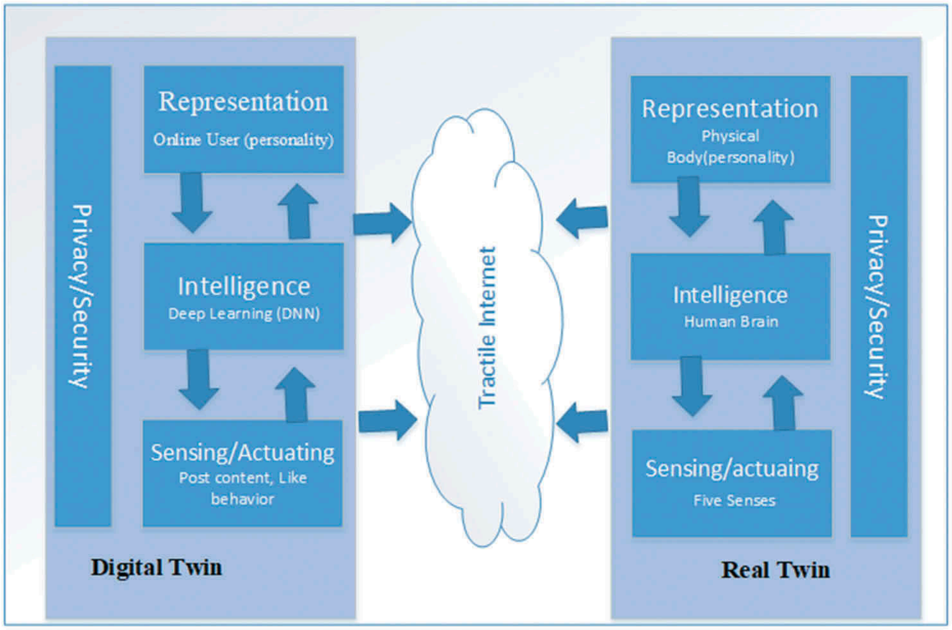

Figure 2. Personality recognition communication framework between digital and real twins.

component without tangible representation depending on the application. In the context of this study, digital twins employ a virtual human representation to understand human personality by timely observing users' posting content and liking behavior.

The words and language text of users' posting content are important means to reflect users' psychology, emotions, and personality. To predict users' Big Five personality accurately, every user's published status texts are aggregated as a user document. A vector representation of a user's document is formed by aggregating the vector of each word as the user's document semantic vector representation. Furthermore, users' online behaviors are recorded and collected by social network service providers. Previous research shows that online behaviors can be employed to identify users' demographic information and personality preferences. Based on this finding, the current work employs users' like behavior records to predict his/her Big Five personality. Likes represent users' positive association (or 'likes') with online content or products, such as photos, books, and web pages. In this study, a user-likes matrix is employed to derive the low-dimensional preference vectors of users. Next, the vectors are incorporated into the deep learning model for personality prediction.

3.1. Text processing

Proper text representation is the basis in the field of text processing. Traditional text processing methods are generally based on the bag-of-words model, which represents a document as a vector of weights for each word. For example, the classic term frequency–inverse document frequency (TFIDF) algorithm represents a document as a vector of TFIDF values for all words by calculating the word frequency and inverse text frequency of each word in a document. Although the bag-of-words model is easy to understand and implement, each word is treated as a separate item, which cannot capture the relationship between words, such as 'king' and 'queen.' The Word2Vec method (Mikolov et al. 2013) is used to solve the above problem. Specifically, the method maps words into a fixed-length vector to express relationships.

First, the status texts posted by each user are aggregated as a document. Next, unrealistic punctuations and stop words are removed. Then, the pretrained Glove-vector (GloVE) file is used to obtain every word vector in the user document. The GloVe model (Pennington, Socher, and Manning 2014) is a new word vector model based on word cooccurrence matrix theory. Based on the statistical word vector and predictive word vector models, the GloVe model employs the word cooccurrence information using the matrix decomposition method. This model pays attention to the context of the Word2Vec window (context) size and leverages the global information. Finally, we can obtain the vector representation of the document through the standardization process.

$$u_D = \frac{1}{n}\sum_{i=1}^{n} w_i, \tag{1}$$

where $w_i \in R^K$ represents the vector (low-dimensional vector) of word i appearing in the user document, K is the dimension of the word vector, n is the total number of words appearing in the article, and $u_D \in R^K$ represents the low-dimensional vector representation of the user document.

3.2. User likes processing

Kosinski, Stillwell, and Graepel (2013) showed that users' online behavior record can be used to accurately estimate their Big Five personality. The authors employed the singular value decomposition (SVD) model to reduce the dimensions of the user-likes matrix. SVD is a type of matrix decomposition and an important feature dimension reduction method in the field of machine learning. Moreover, it is widely used for recommendation systems and natural-language processing in practice. M*N-dimensional matrix A is decomposed by truncated SVD as $A \approx Q_k \sum_k P_k^T$, where $k \ll \min\{M, N\}$. Calculated matrices Q_k and $P_k$contain the k largest eigenvectors of AA^T and A^TA, respectively, thereby achieving dimensionality reduction.

Similarly, the present work utilizes users and their likes to build a user-likes matrix. If a user likes an item, then the corresponding position in the matrix is set to 1, and 0 otherwise. Then, the SVD model is employed to learn the low-dimensional dense representation of the user-likes matrix. Through the SVD model, we can obtain a low-dimensional representation and acquire the like preference vector of each user through the following standardization process.

$$u_L = \frac{1}{m}\sum_{j=1}^{m} l_j, \tag{2}$$

where $l_j \in R^K$ represents the vector (low-dimensional vector) of the user-likes matrix, K is the dimension of the user likes preference vector, m is the total number of items a user likes, and $u_L \in R^K$represents the low-dimensional vector representation of the user likes preference.

3.3. Multitask model for personality prediction

In contrast to general single-task learning in machine learning, multitask learning aims to precomplete different predictive tasks through a model. Multitask learning can utilize useful information between multiple related learning tasks to improve the generalization ability of a model. Moreover, it can be solved by splitting a complex problem (multitasks) into multiple simple and independent subproblems (single tasks). In practical applications, numerous complex problems cannot be split easily, but even if they can, subproblems are interrelated. Multitask learning can solve this problem by sharing data features and model parameters. In image recognition and natural-language processing, multitask learning can improve task performance considerably.

In this work, personality is the combination of human psychological characteristics. Furthermore, personality is a relatively stable organizational structure that can affect people's implicit psychological characteristics and explicit behavior patterns at different times and regions. The Big Five personality describes people's personality through five aspects, namely, openness, conscientiousness, extraversion, agreeableness, and neuroticism. Although the Big Five personality describes users' personality characteristics from five different aspects, these five aspects are closely related in the same user. Every user has a score in the five aspects, and every aspect is indispensable. This type of predictive task is suitable for multitask learning scenarios. To adapt to the multitask learning of Big Five personality prediction, we choose a DNN as the main model. Neural networks mimic the processing mechanisms of the brain and are suitable for predicting complex data patterns. Moreover, compared with other machine learning algorithms, DNNs can effectively capture high-order nonlinear relationships through the superposition of multiple neural layers and activation functions. The model is shown in Figure 3.

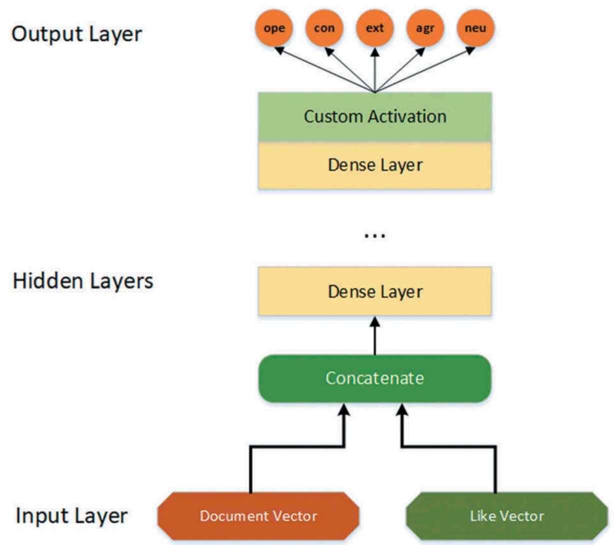

Figure 3. Multitask DNN model.

3.3.1. Input layer

In this model, two types of data input exist, namely, users' document vector and likes vector. The two data combined aggregate in the following way.

$$Input = Concatenate(u_D, u_L) \tag{3}$$

3.3.2. Hidden layers

The hidden layer immediately follows the input layer. The concatenation of multiple hidden layers allows the model to learn deep data interaction. The forward calculation process of the entire model is as follows:

$$\begin{array}{l} h_1 = f_1(W_1 Input + b_1) \\ h_2 = f_2(W_2 h_1 + b_2) \\ \quad\ldots\ldots \\ h_L = f_L(W_L h_{L-1} + b_L) \end{array}, \tag{4}$$

where L represents the Lth hidden layer, and W_l, b_l represents the weight and bias of each neural network layer that must be learned. f_l is an activation function that represents the output of the hidden layer. In practice, we choose the rectifier (ReLU) as the activation function for each layer. Meanwhile, to increase the robustness of the model, a dropout layer behind each hidden layer is added in the model, and the dropout rate is 0.2.

3.3.3. Output layer

This work aims to predict users' Big Five personality simultaneously. Therefore, five outputs are present in the output layer of the model corresponding to users' scores on openness, conscientiousness, extraversion, agreeableness, and neuroticism.

$$\begin{array}{c} \hat{S}_{ui} = f_o(W_o h_L + b_o) \\ \text{with} f_o(x) = 1 + \frac{4}{1+e^{-x}} \end{array}, \tag{5}$$

where $\hat{S}_{ui}$ denotes the user u score on the i-th personality, and $i \in \{openness, conscientiousness, extraversion, agreeableness, neuroticism\}$. W_o, b_o denotes the weight and bias of the output layer. At the output layer, we use a custom activation function to convert the scores of users between 1 and 5.

3.4. Model learning

Unlike general regression tasks, this work focuses on the multitask learning of personality prediction. Therefore, we design a loss function and learn the loss values of five tasks simultaneously.

$$\min L = \frac{1}{|U|} \sum_{u \in U} \sum_{i=1}^{5} (s_{ui} - \hat{s}_{ui})^2 \tag{6}$$

As shown above, this model uses the mean square error as the objective function to train the model, where u denotes the user, U represents the set of users, $|U|$is the number of users, $\hat{s}_{ui}$ is the score of user u on the i-th personality predicted by the model, and s_{ui} is the corresponding real value.

The model is implemented using the open source Keras,[1] and the parameters are learned using Adam. The training process is run for 50 iterations. The detailed learning algorithm is shown in Algorithm 1.

Algorithm 1. Learning algorithm of multitask model

Input: Document vector matrix $R_D \in \mathbb{R}^{|U| \times K}$, each row of the matrix represents a user's document vector u_D, likes vector matrix $R_L \in \mathbb{R}^{|U| \times K}$, each row of the matrix represents a user's likes vector u_L, users' Big Five personality matrix $R_p \in \mathbb{R}^{|U| \times 5}$, each row of the matrix represents a user's five scores on the Big Five personality, batch size m; max epoch T

Output:$\Theta = [\mathrm{W}, \mathrm{b}]$ the weight $w_{ij} \in \mathrm{W}$ and bias $b_j \in \mathrm{b}$ in the model

1: Initialize Θ with a Gaussian distribution with a mean of 0 and a standard variation of 0.1

2: For epoch $\leftarrow$ 1 to T **do**

3: Obtain training data D

4: For mini epoch $\leftarrow$ 1 to $\frac{|D|}{m}$ do

5: Compute the loss L

6: Calculate the partial derivative of the loss function versus the parameter $\frac{\partial L}{\partial w_{ij}}, \frac{\partial L}{\partial b_j}$

7: Update Θ with learn rate

8: **end for**

9: **end for**

10: Return $\Theta = [W, b]$

4. Experiments and results

4.1. Experimental settings

4.1.1. Dataset

The experiments employ a portion of the data from the *MyPersonality* project.[2] The *MyPersonality* project records the Big Five personality scores published by Facebook users and rich network behavior data, including users' friend relationships, groups, published status content, and favorites record. In the final dataset, we sample 22,000 users, 3,183,816 corresponding post statuses, and 301,105 likes. The statistics of the dataset are shown in Table 1. Meanwhile, the statistics of the Big Five personality scores for all users are shown in Figure 4. In the experiment, we randomly divide the training and testing sets by 80%/20%. Moreover, to better tune the parameters, we randomly select 10% from the training set as the validation set.

Table 1. Statistics of dataset.

Item	Number
User	22,000
Status	3,183,816
Status numbers/user	144.7189091
Liked items	301,105
Likes record	3,864,409
Likes record/user	175.6549545

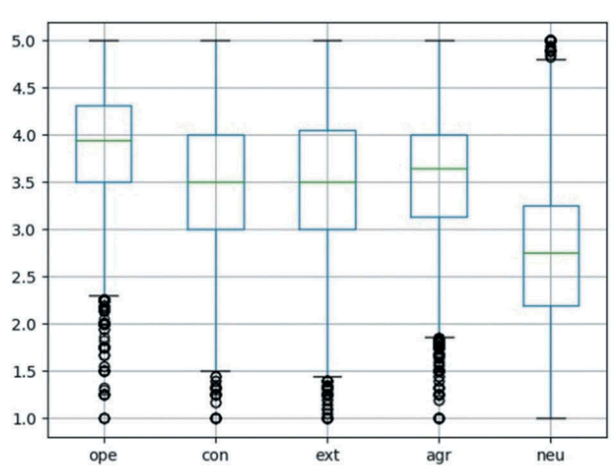

Figure 4. Statistics of users' Big Five personality scores.

4.1.2. Evaluation metrics

In this work, we use mean absolute error (MAE) to evaluate the proposed method. MAE is widely used in personality prediction tasks, where n represents the number of samples in the test set, $\hat{s}_{ui}$ is the predicted score of user u on the i-th personality, and s_{ui} is the true value. Given that MAE is an error calculation, the lower the MAE, the stronger the learning ability of the model.

$$MAE = \frac{1}{n}\sum_{u \in U}\sum_{i=1}^{5}|s_{ui} - \hat{s}_{ui}| \tag{7}$$

4.1.3. Parameter setting

Two hyperparameters can have a large impact on the experimental results. The first parameter is the dimension of the user document vector, and the other parameter is the dimension in the SVD model. In general, the larger the dimension, the more the information that can be represented and the better the model. To verify whether this intuition is correct, several comparative experiments are conducted to verify the effect of different dimensions on the results. The three dimensions provided by the GloVE model are divided into 100, 200, and 300. Therefore, we define the output dimension of SVD as 100, 200, and 300. The results of the experiments are shown in Table 2. In Table 2, the first column (Feature) indicates the features used in the model. The abbreviation 'DV' indicates the document vector extracted in Section 3.1, which is the users' text feature. The abbreviation 'LV' denotes the user's likes vector obtained from Section 3.2, which is the users' liking behavior feature.

As shown in Table 2, using high-dimensional vectors is useful for improving prediction results when using only the status text published by the users. However, high-dimensional vectors are not as effective as low-dimensional vectors for the prediction of extraversion and agreeableness. In the next study, the experiments use the 300-dimensional text vector to predict the Big Five personality. In contrast to text data, the vectors of the user-likes matrix built with

Table 2. MAE obtained from DNN model with different dimensions.

Feature	Dimension	ope	con	ext	agr	neu
DV	100	0.524727	0.565691	0.637045	**0.559145**	0.650402
DV	200	0.521419	0.565946	**0.634641**	0.560098	0.642806
DV	300	**0.519473**	**0.562108**	0.636015	0.563406	**0.640859**
LV	100	**0.510742**	0.568343	**0.633011**	**0.559099**	**0.638539**
LV	200	0.513919	**0.567201**	0.635606	0.559568	0.638833
LV	300	0.520375	0.573119	0.63786	0.560847	0.643589

Table 3. MAE obtained from DNN model with different network structures.

Feature	Network structure	Ope	Con	Ext	Agr	Neu
DV + LV	[500]	0.506718	0.560451	0.627927	0.564151	0.63874
DV + LV	[500, 200]	0.505175	0.559196	0.626359	**0.553663**	0.637454
DV + LV	[500, 300, 200]	0.504031	0.559348	0.623858	0.554446	0.635637
DV + LV	[500, 300, 200, 100]	**0.503842**	**0.558871**	**0.6232**	0.553688	**0.63543**
DV + LV	[500, 400, 300, 200, 100]	0.504747	0.559839	0.628215	0.564388	0.638926

SVD perform well in low-dimensional situations. The reason may be because users and items are not particularly large; hence low-dimensional space is adequate to capture valid information.

Another important hyperparameter that can affect the experimental results is the number of layers in the DNN and the number of neurons in each layer. To verify whether the deep structure has a significant effect on the experimental results, we test different neural layers with different numbers of neurons. The experimental results are shown in Table 3.

The second column of Table 3 is the network structure. For example, [500, 300, 200] indicates that the deep model has three layers, in which the number of neurons in the first layer is 500, and the number of neurons in the second layer is 300. Moreover, the number of neurons in the third layer is 200. As the number of layers in the neural network model increases, the parameters of the model become considerably complex. In addition, learning ability is improved but not as deep as possible. Table 3 demonstrates that the five-layer neural network model is not better than the four-layer model. Therefore, the four-layer network structure is employed in the model. Combined with Table 2, models that aggregate text and user-likes features perform better than models that use only one type of data.

4.2. Method performance comparison

To compare the models, we select the LR and support vector regression (SVR) models as baselines. These two algorithms are widely used in regression problems. LR models can predict personality by combining different features linearly. SVR models employ SVMs to fit data, find regression planes, and perform regression analysis. To compare the features, we choose traditional TFIDF features as baseline features (widely used in text analysis) and compare them with the proposed method.

Table 4. Performance comparison of methods.

Feature	Model	Ope	Con	Ext	Agr	Neu
TFIDF	LR	0.601265	0.645289	0.713144	0.64739	0.741121
TFIDF	SVR	0.54854	0.593799	0.665518	0.569239	0.652625
DV	LR	0.523068	0.56697	0.643654	0.561265	0.64659
DV	SVR	0.535418	0.576882	0.654423	0.561301	0.646419
DV	DNN	0.519473	0.562108	0.636015	0.563406	0.640859
DV + LV	DNN	**0.503842**	**0.558871**	**0.6232**	**0.553688**	**0.63543**

Table 4 indicates that whether it is an LR or SVR model, performance when using TFIDF features is not as good as performance when using the extracted document vector. This finding proves that word representation in vector form can better obtain semantic information for further analysis. Furthermore, when using document vectors as features, the DNN used in this work is better than the LR and SVR models. This finding shows that the DNN is better than the general machine learning algorithm in this problem. Finally, the best combination of document and likes vectors obtained with SVD performs best. Users' likes record is conducive to predict users' Big Five personality, and the fusion of multiple behavioral data is superior for inferring users' personality characteristics.

4.3. Further analysis

Further analysis is conducted to explore comprehensively the effect of posting status and liking behavior

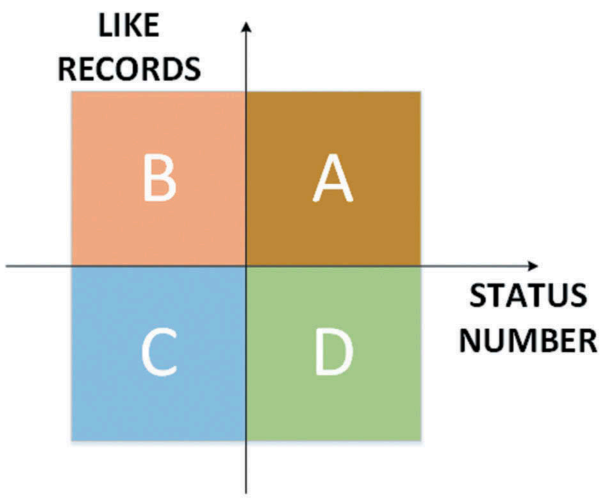

Figure 5. User division diagram.

on personality prediction. The median numbers of users' postings (96) and likes record (73) are calculated to divide the users of the test set into four categories, as shown in Figure 5.

The average MAE results for the four different categories of users are presented in Table 5.

Table 5 illustrates that A-type users' Big Five personality has the highest prediction accuracy (smallest MAE), and C-type users have the lowest accuracy (maximum MAE). These results signify that if a user is active on the Internet (with much status or likes record), then we can likely infer the user's personality traits. This finding explains the main reason we use rich network behavior data to predict user personality.

Meanwhile, A-type and B-type users have similar MAE values, whereas C-type and D-type users have similar MAE values (except ope), and the MAEs of A/B users are generally better than those of C/D users. Increased posting status data do not enhance personality prediction performance. However, increased user likes record enhances the model's ability to predict personality. Compared with text data, users' likes record is more useful for deriving users' personality traits. A possible reason could be, when users publish their status, they must consider how to organize the text and its influence on their audience. In users' likes record, they only 'like' things they find pleasing, thereby reflecting their characteristics.

4.4. Implications in the manufacturing industry

Companies pay attention to the mode of personalized production, which is driven by customer needs. To meet the mass personalized requirements of customers, firms employ various types of innovation modes, from traditional inventory sales modes to customer reversed customization models. Newly generated information technologies and online social networks provide convenient and low-cost communication platforms for consumers and manufacturers. Consumers purchase products increasingly in pursuit of personalization and experience. Personality prediction methods are useful to identify consumer preferences and guide the personalized manufacturing process. Industry manufacturing is evolving from automation to intelligence. Meanwhile, smart manufacturing and digital twin applications should consider complete data generated from the entire manufacturing cycle. Specifically, data generated by consumers can reflect the advantages and disadvantages of products directly and play an important role in improving product development and production processes.

The production and development of manufacturing products aim to meet the needs of consumers. However, the needs of different consumers differ and are continuously changing. Therefore, the user profile for a product can help the manufacturing industry accurately and easily understand the market positioning of a product and the market segmentation of the target consumer. This work predicts the personality of target consumers through user-generated behavioral data. This prediction can effectively help the manufacturing industry understand target consumers, thereby improving product design and the manufacturing process. Given that the generation of consumer data is dynamic, this work can track changes in target

Table 5. Average MAE of different category users.

Category	User numbers	Ope	Con	Ext	Agr	Neu
A	1,353	0.481569	0.552464	0.61563	0.530382	0.628472
B	788	0.482564	0.527393	0.612575	0.546871	0.634783
C	1,422	0.550607	0.578554	0.643568	0.583681	0.665056
D	837	0.507368	0.576618	0.654707	0.575215	0.636716

consumers and help companies design marketing strategies for different products.

5. Conclusions and future work

In this research, we focus on the digital twin issue to understand individuals. A deep learning-based method is proposed to predict individuals' personality traits. Status content published by users and their liking behavior record are employed to predict users' Big Five personality. When preprocessing data, the text data are transformed into the users' document vector through the pretrained word vector model. Meanwhile, liking behavior data are transformed into the users' preference vector through SVD techniques. The two vectors are aggregated and inputted into the multitask DNN model to calculate the users' Big Five personality. The experiments show that the proposed approach is more effective than traditional methods. Meanwhile, the experiments demonstrate that compared with text data, liking behavior data are more useful for predicting users' personality traits.

Predicting users' personality is a very challenging and practical task. Future works should preprocess data accurately. For example, the traditional text preprocessing process may not be suitable for Big Five personality prediction. Certain stop words and punctuation marks play an important role in prediction tasks. In addition, this work does not consider the order relationship of words. In future research, we can use the long short-term memory model and attention mechanism to model text accurately. Meanwhile, through digital twins, we hope to integrate data from different spaces, such as manufacturing experts' knowledge and consumers' purchase feedback data, to accurately guide companies' production planning and product design.

Notes

1. https://keras.io.
2. http://mypersonality.org.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work is supported by the National Natural Science Foundation of China (71872060, 91846201, 71490725, 71722010, 91546114, and 91746302), 2019 Guangdong Special Support Talent Program – Innovation and Entrepreneurship Leading Team (China) (2019BT02S593) and the Fundamental Research Funds for the Central Universities of China (JZ2020HGPA0113), partially Sponsored by Zhejiang Lab (No. 2019KE0AB04).

ORCID

Zhiqiang Tian http://orcid.org/0000-0002-2296-4543
Yelin Fu http://orcid.org/0000-0003-4515-2211

References

Aivaliotis, P., K. Georgoulias, and G. Chryssolouris. 2019. "The Use of Digital Twin for Predictive Maintenance in Manufacturing." *International Journal of Computer Integrated Manufacturing* 32 (11): 1067–1080. doi:10.1080/0951192X.2019.1686173.

Amichai-Hamburger, Y., and G. Vinitzky. 2010. "Social Network Use and Personality." *Computers in Human Behavior* 26 (6): 1289–1295. doi:10.1016/j.chb.2010.03.018.

Back, M. D., J. M. Stopfer, S. Vazire, S. Gaddis, S. C. Schmukle, B. Egloff, and S. D. Gosling. 2010. "Facebook Profiles Reflect Actual Personality, Not Self-idealization." *Psychological Science* 21 (3): 372–374. doi:10.1177/0956797609360756.

Baik, J., K. Lee, S. Lee, Y. Kim, and J. Choi. 2016. "Predicting Personality Traits Related to Consumer Behavior Using SNS Analysis." *New Review of Hypermedia and Multimedia* 22 (3): 189–206. doi:10.1080/13614568.2016.1152313.

Banerjee, A., R. Dalal, S. Mittal, and K. P. Joshi. 2017. "Generating Digital Twin Models Using Knowledge Graphs for Industrial Production Lines." *UMBC Information Systems Department*. doi:10.1145/3091478.3162383.

Bielefeldt, B., J. Hochhalter, and D. Hartl. 2015. "Computationally Efficient Analysis of SMA Sensory Particles Embedded in Complex Aerostructures Using a Substructure Approach." ASME 2015 Conference on Smart Materials, Adaptive Structures and Intelligent Systems. doi: 10.1115/SMASIS2015-8975.

Blignaut, P., and N. Annelie. 2008. "The Influence of Temperament Style on a Student's Choice of and Performance in a Computer Programming Course." *Computers in Human Behavior* 24 (3): 1010–1020. doi:10.1016/j.chb.2007.03.005.

Bruynseels, K., F. Santoni de Sio, and J. Hoven. 2018. "Digital Twins in Health Care: Ethical Implications of an Emerging Engineering Paradigm." *Frontiers in Genetics* 9: 31. doi:10.3389/fgene.2018.00031.

Chen, T. Y., M. C. Tsai, and Y. M. Chen. 2016. “A User's Personality Prediction Approach by Mining Network Interaction Behaviors on Facebook.” *Online Information Review* 40 (7): 913–937. doi:10.1108/OIR-08-2015-0267.

Collobert, R., and J. Weston. 2008. “A Unified Architecture for Natural Language Processing: Deep Neural Networks with Multitask Learning.” Proceedings of the 25th international conference on Machine learning. doi: 10.1145/1390156.1390177.

El Saddik, A. 2018. “Digital Twins: The Convergence of Multimedia Technologies.” *IEEE MultiMedia* 25 (2): 87–92. doi:10.1109/MMUL.2018.023121167.

Farnadi, G., G. Sitaraman, S. Sushmita, F. Celli, M. Kosinski, D. Stillwell, S. Davalos, M. F. Moens, and M. D. Cock. 2016. “Computational Personality Recognition in Social Media.” *User Modeling and User-adapted Interaction* 26 (2–3): 109–142. doi:10.1007/s11257-016-9171-0.

Farnadi, G., S. Zoghbi, M. F. Moens, and M. D. Cock. 2013. “Recognising Personality Traits Using Facebook Status Updates.” Seventh International AAAI Conference on Weblogs and Social Media, Cambridge, Massachusetts, USA.

Funder, D. C. 2001. “Personality.” *Annual Review of Psychology* 52 (1): 197–221. doi:10.1146/annurev.psych.52.1.197.

Gavrilescu, M., and N. Vizireanu. 2017. “Predicting the Sixteen Personality Factors (16PF) of an Individual by Analyzing Facial Features.” *EURASIP Journal on Image and Video Processing* 2017 (1): 59. doi:10.1186/s13640-017-0211-4.

Golbeck, J., C. Robles, and K. Turner. 2011. “Predicting Personality with Social Media.” CHI'11 extended abstracts on human factors in computing systems. doi: 10.1145/1979742.1979614.

Hassanein, M., W. Hussein, S. Rady, and T. F. Gharib. 2018. “Predicting Personality Traits from Social Media Using Text Semantics.” 2018 13th International Conference on Computer Engineering and Systems (ICCES). doi: 10.1109/ICCES.2018.8639408.

Hearst, M. A., S. T. Dumais, E. Osuna, J. Platt, and B. Scholkopf. 1998. “Support Vector Machines.” *IEEE Intelligent Systems and Their Applications* 13 (4): 18–28. doi:10.1109/5254.708428.

Hochhalter, J., W. P. Leser, J. A. Newman, V. K. Gupta, V. Yamakov, S. R. Cornell, and G. Heber. 2014. “Coupling Damage-Sensing Particles to the Digitial Twin Concept.” Langley Research Center, National Aeronautics and Space Administration, Hampton, VA. NASA/TM-2014-218257, L-20401, NF1676L-18764.

Judge, T. A., C. A. Higgins, C. J. Thoresen, and M. R. Barrick. 1999. “The Big Five Personality Traits, General Mental Ability, and Career Success across the Life Span.” *Personnel Psychology* 52 (3): 621–652. doi:10.1111/j.1744-6570.1999.tb00174.x.

Knapp, M. L., and J. A. Daly. 2002. *Handbook of Interpersonal Communication*. London, United Kingdom: Sage.

Kosinski, M., D. Stillwell, and T. Graepel. 2013. “Private Traits and Attributes are Predictable from Digital Records of Human Behavior.” *Proceedings of the National Academy of Sciences* 110 (15): 5802–5805. doi:10.1073/pnas.1218772110.

Li, J., F. Tao, Y. Cheng, and L. Zhao. 2015. “Big Data in Product Lifecycle Management.” *The International Journal of Advanced Manufacturing Technology* 81 (1–4): 667–684. doi:10.1007/s00170-015-7151-x.

Li, Y., T. Zhu, A. Li, F. Zhang, and X. Xu. 2011. “Web Behavior and Personality: A Review.” 2011 3rd Symposium on Web Society. doi: 10.1109/SWS.2011.6101275.

Liu, Y., J. Wang, and Y. Jiang. 2016. “PT-LDA: A Latent Variable Model to Predict Personality Traits of Social Network Users.” *Neurocomputing* 210: 155–163. doi:10.1016/j.neucom.2015.10.144.

Mairesse, F., M. A. Walker, M. R. Mehl, and R. K. Moore. 2007. “Using Linguistic Cues for the Automatic Recognition of Personality in Conversation and Text.” *Journal of Artificial Intelligence Research* 30: 457–500. doi:10.1613/jair.2349.

Marston, M. W. 2013. *Emotions of Normal People*. Routledge. doi:10.1037/13390-000.

McCrae, R. R., and P. T. C. Jr. 1997. “Personality Trait Structure as a Human Universal.” *American Psychologist* 52 (5): 509. doi:10.1037//0003-066X.52.5.509.

Mikolov, T., I. Sutskever, K. Chen, G. S. Corrado, and J. Dean. 2013. “Distributed Representations of Words and Phrases and Their Compositionality.” Advances in neural information processing systems, Lake Tahoe, Nevada, USA.

Moffitt, K., J. Giboney, E. Ehrhardt, J. K. Burgoon, and J. F. Nunamaker. 2010. “Structured Programming for Linguistic Cue Extraction.” *The Center for the Management of Information*. http://splice.cmi.arizona.edu

Mohammad, S., X. Zhu, and J. Martin. 2014. “Semantic Role Labeling of Emotions in Tweets.” Proceedings of the 5th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, Baltimore, Maryland, USA.

Mønsted, B., A. Mollgaard, and J. Mathiesen. 2018. “Phone-based Metric as a Predictor for Basic Personality Traits.” *Journal of Research in Personality* 74: 16–22. doi:10.1016/j.jrp.2017.12.004.

Myers, I. B., M. H. McCaulley, N. L. Quenk, and A. L. Hammer. 2003. *MBTI Manual: A Guide to the Development and Use of the Myers-Briggs Type Indicator, 3rd*. Palo Alto, CA: Consulting Psychologists Press.

Nave, G., J. Minxha, D. M. Greenberg, M. Kosinski, D. Stillwell, and J. Rentfrow. 2018. “Musical Preferences Predict Personality: Evidence from Active Listening and Facebook Likes.” *Psychological Science* 29 (7): 1145–1158. doi:10.1177/0956797618761659.

Ozer, D. J., and V. Benet-Martinez. 2006. “Personality and the Prediction of Consequential Outcomes.” *Annual Review of Psychology* 57: 401–421. doi:10.1146/annurev.psych.57.102904.190127.

Park, G., H. A. Schwartz, J. C. Eichstaedt, M. L. Kern, M. Kosinski, D. J. Stillwell, L. H. Ungar, and M. E. P. Seligman. 2015. “Automatic Personality Assessment through Social Media Language.” *Journal of Personality and Social Psychology* 108 (6): 934. doi:10.1037/pspp0000020.

Park, K. T., Y. W. Nam, H. S. Lee, S. J. Im, S. D. Noh, J. Y. Son, and H. Kim. 2019. “Design and Implementation of a Digital Twin Application for a Connected Micro Smart Factory.”

International Journal of Computer Integrated Manufacturing 32 (6): 596–614. doi:10.1080/0951192X.2019.1599439.

Pennebaker, J. W., and L. A. King. 1999. “Linguistic Styles: Language Use as an Individual Difference.” *Journal of Personality and Social Psychology* 77 (6): 1296. doi:10.1037//0022-3514.77.6.1296.

Pennington, J., R. Socher, and C. Manning. 2014. “Glove: Global Vectors for Word Representation.” Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP), Doha, Qatar.

Ross, C., E. S. Orr, M. Sisic, J. M. Arseneault, M. G. Simmering, and R. R. Orr. 2009. “Personality and Motivations Associated with Facebook Use.” *Computers in Human Behavior* 25 (2): 578–586. doi:10.1016/j.chb.2008.12.024.

Samani, Z. R., S. C. Guntuku, M. E. Moghaddam, D. Preoţiuc-Pietro, and L. H. Ungar. 2018. “Cross-platform and Cross-interaction Study of User Personality Based on Images on Twitter and Flickr.” *PloS One* 13 (7): e0198660. doi:10.1371/journal.pone.0198660.

Schleich, B., N. Anwer, L. Mathieu, and S. Wartzack. 2017. “Shaping the Digital Twin for Design and Production Engineering.” *CIRP Annals* 66 (1): 141–144. doi:10.1016/j.cirp.2017.04.040.

Smarslok, B., A. Culler, and S. Mahadevan. 2012. “Error Quantification and Confidence Assessment of Aerothermal Model Predictions for Hypersonic Aircraft.” 53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference 20th AIAA/ASME/AHS Adaptive Structures Conference 14th AIAA. doi: 10.2514/6.2012-1817.

Söderberg, R., K. Wärmefjord, J. S. Carlson, and L. Lindkvist. 2017. “Toward a Digital Twin for Real-time Geometry Assurance in Individualized Production.” *CIRP Annals* 66 (1): 137–140. doi:10.1016/j.cirp.2017.04.038.

Stark, R., S. Kind, and S. Neumeyer. 2017. “Innovations in Digital Modelling for Next Generation Manufacturing System Design.” *CIRP Annals* 66 (1): 169–172. doi:10.1016/j.cirp.2017.04.045.

Tadesse, M. M., H. Lin, B. Xu, and L. Yang. 2018. “Personality Predictions Based on User Behavior on the Facebook Social Media Platform.” *IEEE Access* 6: 61959–61969. doi:10.1109/ACCESS.2018.2876502.

Tao, F., F. Sui, A. Liu, Q. Qi, M. Zhang, B. Song, Z. Guo, S. C.-Y. Lu, and A. Y. C. Nee. 2018. “Digital Twin-driven Product Design Framework.” *International Journal of Production Research* 1–19. doi:10.1080/00207543.2018.1443229.

Tao, F., F. Sui, A. Liu, Q. Qi, M. Zhang, B. Song, and A. Y. C. Nee. 2019. “Digital Twin-driven Product Design Framework.” *International Journal of Production Research* 57 (12): 3935–3953. doi:10.1080/00207543.2018.1443229.

Tao, F., H. Zhang, A. Liu., and A. Y. Nee. 2018b. “Digital Twin in Industry: State-of-the-art.” *IEEE Transactions on Industrial Informatics* 15 (4): 2405–2415. doi:10.1109/TII.2018.2873186.

Tao, F., M. Zhang, Y. Liu, and A. Y. C. Nee. 2018a. “Digital Twin Driven Prognostics and Health Management for Complex Equipment.” *CIRP Annals* 67 (1): 169–172. doi:10.1016/j.cirp.2018.04.055.

Tuegel, E. J., A. R. Ingraffea, T. G. Eason, and S. M. Spottswood. 2011. “Reengineering Aircraft Structural Life Prediction Using a Digital Twin.” *International Journal of Aerospace Engineering* 2011. doi:10.1155/2011/154798.

Wu, Y., M. Kosinski, and D. Stillwell. 2015. “Computer-based Personality Judgments are More Accurate than Those Made by Humans.” *Proceedings of the National Academy of Sciences* 112 (4): 1036–1040. doi:10.1073/pnas.1418680112.

Xue, D., L. F. Wu, Z. Hong, S. Guo, L. Gao, Z. Wu, X. Zhong, and J. Sun. 2018. “Deep Learning-based Personality Recognition from Text Posts of Online Social Networks.” *Applied Intelligence* 48 (11): 4232–4246. doi:10.1007/s10489-018-1212-4.

Zhang, C., W. Xu, J. Liu, Z. Liu, Z. Zhou, and D. T. Pham. 2019. “Digital Twin-enabled Reconfigurable Modeling for Smart Manufacturing Systems.” *International Journal of Computer Integrated Manufacturing* 1–25. doi:10.1080/0951192X.2019.1699256.

Zhang, H., Q. Liu, X. Chen, D. Zhang, and J. Leng. 2017a. “A Digital Twin-based Approach for Designing and Multi-objective Optimization of Hollow Glass Production Line.” *Ieee Access* 5: 26901–26911. doi:10.1109/ACCESS.2017.2766453.

Zhang, T., R. Z. Qin, Q. L. Dong, W. Gao, H. R. Xu, and Z. Y. Hu. 2017b. “Physiognomy: Personality Traits Prediction by Learning.” *International Journal of Automation and Computing* 14 (4): 386–395. doi:10.1007/s11633-017-1085-8.

Zhang, W., R. Li, T. Zeng, Q. Sun, S. Kumar, J. Ye, and S. Ji. 2016. “Deep Model Based Transfer and Multi-task Learning for Biological Image Analysis.” *IEEE Transactions on Big Data* 1. doi:10.1145/2783258.2783304.