Extracting User Preferences and Personality from Text for Restaurant Recommendation

Evripides Christodoulou
ep.xristodoulou@edu.cut.ac.cy
Cyprus University of Technology
Limassol, Cyprus

Andreas Gregoriades
andreas.gregoriades@cut.ac.cy
Cyprus University of Technology
Limassol, Cyprus

Herodotos Herodotou
herodotos.herodotou@cut.ac.cy
Cyprus University of Technology
Limassol, Cyprus

Maria Pampaka
maria.pampaka@manchester.ac.uk
The University of Manchester
Manchester, UK

ABSTRACT

Restaurant recommender systems are designed to support restaurant selection by assisting consumers with the information overload problem. However, despite their promises, they have been criticized of insufficient performance. Recent research in recommender systems has acknowledged the importance of personality in improving recommendation; however, limited work exploited this aspect in the restaurant domain. Similarly, the importance of user preferences in food has been known to improve recommendation but most systems explicitly ask the users for this information. In this paper, we explore the influence of personality and user preference by utilizing text in consumers' electronic word of mouth (eWOM) to predict the probability of a user enjoying a restaurant he/she had not visited before. Food preferences are extracted though a trained named-entity recognizer learned from a labelled dataset of foods, generated using a rule-based approach. The prediction of user personality is achieved through a bi-directional transformer approach with a feed-forward classification layer, due to its improved performance in similar problems over other machine learning models. The personality classification model utilizes the textual information of reviews and predicts the personality of the author. Topic modelling is used to identify additional features that characterize users' preferences and restaurants properties. All aforementioned features are used collectively to train an extreme gradient boosting tree model, which outputs the predicted user rating of restaurants. The trained model is compared against popular recommendation techniques such as nonnegative matrix factorization and single value decomposition.

CCS CONCEPTS

• **Information systems → Recommender systems**; **Information extraction**.

KEYWORDS

Consumer Personality, Food preference extraction, Recommender System, Topic Modelling

ORSUM@ACM RecSys 2022, September 23rd, 2022, Seattle, WA, USA

Reference Format:
Evripides Christodoulou, Andreas Gregoriades, Herodotos Herodotou, and Maria Pampaka. 2022. Extracting User Preferences and Personality from Text for Restaurant Recommendation. In *5th Workshop on Online Recommender Systems and User Modeling (ORSUM 2022), in conjunction with the 16th ACM Conference on Recommender Systems, September 23rd, 2022, Seattle, WA, USA.*

1 INTRODUCTION

Dining is one of the top five tourist activities during a leisure trip that plays a central role in travel experience. Recently, interest on food experience has been growing [18], with businesses in the hospitality sector seeking insights regarding the dining behaviors and preferences of customers to improve decision making in areas such as marketing [23] and recommendation [2]. Past research that utilizes food in recommender systems such as [11] employ simple techniques such as frequencies of food vocabularies in Bag of Words. However, such techniques require a lexicon of complete list of foods that usually is not available for different cousins and countries. In this paper, we utilize implicit and explicit information of consumers' eWOM to improve restaurant recommendation. Implicit information refers to textual comments in reviews that can be used to estimate consumers' personality and preferences, while explicit features refer to ratings of restaurants, their estimated value, price, and cuisine offered.

There are several ways to extract user preferences for restaurant recommendation [4]. The simplest is through explicit queries by asking users to define their preferences. This however has some disadvantages, as food preferences might not be covered by the questions asked. Alternative methods utilize user ratings to find similarities between users and restaurants (e.g., collaborative filtering). Another method for preference extraction is user opinion analysis that utilizes natural language processing.

Traditional recommendation approaches base their recommendations on user preferences extracted from users' historical records, such as ratings, reviews, or purchases. Popular techniques include the collaborative and content-based filtering approaches. Recently, there is strong interest in the utilization of users' personality, since it is linked to perception, motivation, and preference, and is known to remain stable during adulthood. Personality is directly associated with consumer emotions and has strong impact on satisfaction with theory indicating that people with the same personality have similar preferences and needs [9]. The application of users' personality has enhanced the performance of recommender systems in the tourism domain with results improving point of interest, destination recommendations, utilizing either questionnaires or automated personality recognition.

This paper illustrates the utilization of consumers' personality and user preferences extracted from the textual part of electronic word of mouth (eWOM) to improve recommendation. EWOM represent consumer opinions about products and services and has been used extensively in identifying consumers' preferences. Recommendations are made by training an Extreme Gradient Boosting (XGBoost) prediction model using as features, the users' personality and the users' preferences (e.g., food). XGBoost is used due to its good performance in similar recommendation problems [32]. The research question addressed in this paper focuses on whether the integration of personality with other features inferred from structured and unstructured parts of online reviews improves restaurant recommendation, in contrast to popular model-based collaborative filtering (CF) techniques such as nonnegative matrix factorization (NMF) and single value decomposition (SVD).

The proposed approach utilizes consumers' food preferences and personalities along with perceptions about venues from eWOM to recommend most suitable restaurants to tourists. Labelled personality data is utilized to train a BERT (Bidirectional Encoder Representations from Transformers) classifier using the personality model of Myers-Briggs Type Indicator (MBTI) due to its good results in previous studies [6]. User preferences are extracted through topic modelling and a trained food named entity recognizer. An XGBoost model is generated to predict the probability of a user liking an unvisited restaurant based on its personality, preference, and themes that characterize the venue.

The research question addressed in this work is how to best combine user preference and personality models with topic features inferred from eWOM to produce the best recommendation, in contrast to popular model-based collaborative filtering (CF) techniques. This is a continuation of our previous work in [7, 10] that examine the use of personality and emotion in recommender systems. The contribution of this work lies in the automated detection of food preferences from eWOM and its combination with user personality and topic modeling for restaurant recommendation.

The paper is organized as follows. The next section introduces background knowledge on restaurant recommender systems and techniques for extracting food preferences and personality from text. The next section describes techniques for identifying topics discussed in consumers' eWOM and personality prediction using deep neural networks. Subsequent sections elaborate on the methodology followed and the results obtained. The paper concludes with the discussion and future directions.

2 EXISTING KNOWLEDGE

This section provides a review of recommendation techniques, the concept of personality, and elaborates on how it has been used in recommender systems so far.

2.1 Restaurant Recommender Systems

Recommender systems aim to predict the satisfaction of a consumer with an item (product/service) he/she has not bought yet [22]. This is part of one-to-one marketing that seeks to match items to consumers' preferences in contrast to mass marketing aiming to satisfy a target market segment [3]. Popular approaches focus on consumers' past experiences (ratings) for the creation of a user-item matrix and based on that predict what is more appropriate to a user depending on either similarity between users or items (products, services) [22]. The relationship between consumers or between products can be found using similarity metrics, and this method is known as Collaborative Filtering (CF) [1]. This has been successfully applied in tourism recommendation problems such as hotels or points of interests, and is considered as one of the most popular techniques [26]. Another popular technique is content-based filtering, that attempts to guess what a user may like based on items' features rather than their rating [1]. A hybrid approach takes the advantage of both content-based filtering and collaborative filtering [22].

CF techniques, however, suffer from the cold start problem that occurs when very little or no data is available about a user and thus inability to identify similar consumers [33]. In addition, data sparsity exacerbates the problem when there are a lot of unrated items in the user-item matrix. This occurs when there is not enough data to populate the user-item matrix based on which to make reliable inferences [24]. In tourism, the collection of data is difficult and time-consuming due to the limited time that tourists spent at a destination. The cold start problem appears with first-time users (tourists) since there are no records of their purchasing activity at a specific destination.To address these CF problems, recent methods utilise machine learning techniques such as matrix factorisation to approximate the user-item matrix content using latent variables that emerge from the initial data. The singular value decomposition (SVD), optimized SVD (SVD++), and non-negative matrix factorization (NMF) models factorize the user-item matrix and predict the satisfaction of users for products that are unknown [19]. Alternatively, content-based approaches utilize metadata about new products to address the cold start problem. A useful source for obtaining these metadata is textual information from eWOM and its analysis using text analytics [8]. An example application includes work by Sun et al. [37] that improved CF performance by analysing restaurants eWOM to define numerical features corresponding to consumers satisfaction through sentiment analysis. In the same vein, topic modelling techniques have been used with CF to assist in estimating the similarity between consumers or items [12]. Finally, work by Zhang et al. [41] used consumers or items characteristics to cluster them into groups, and then find correlations between clusters to address the data sparsity problem.

Recently, a strong interest emerged in using the personalities of consumers in an effort to better understand and match their needs, as "personality" relates to the perceptions, feelings, motivations, and preferences of individuals [35]. The application of user personality has improved the performance of recommendations in the tourism domain for points of interest compared to traditional methods [38]. Personality-based recommendations have also been shown to greatly reduce the cold start and data sparsity problems, and improved the performance of recommendations in areas such as online advertising, social media, books, and music [39]. However, these approaches do not take advantage of eWOM data from users on the web to extract their preferences and their personalities. They focus mainly on the extraction of user data from specialized questionnaires to collect consumers' behaviours and personalities. Such approaches fail to continuously update the system because of the time-consuming use of questionnaires that leads to the loss of automation and update limitations.

2.2 Personality Extraction from Text

Personality is a set of characteristics and behaviours of an individual that influence many areas of his/her life such as motivations, preferences, as well as consumer preferences and behaviour [38]. Applications of automated personality predictions have been applied by researchers on data from various social networks such as Facebook, Twitter, to explore correlations between personalities and the different user activities, purchasing behaviors and liking of foods from specific cuisines [16].

The two most popular text-based personality classification methods are based on the Myers-Briggs Type Indicator (MBTI) [5] and the Big Five [31] personality traits due to the availability of labelled data on these models. The classifiers with best performance are usually employing the MBTI personality model that focuses on 8 key types of characteristics that people have, Extraversion or Introversion, Sensing or Intuition, Thinking or Feeling, and Judging or Perceiving, behaviours. The combination of characteristics can shape 16 different personality types and classify people to the proper personality cluster [1]. The Big 5 Personality model express personality in the following 5 dimensions: Agreeableness, Extraversion, Openness to Experience, Conscientiousness, and Neuroticism. Such taxonomies are recognized as a valid mechanism for defining the most essential aspects of personality that describes people characteristics that creates and reflects their behaviour [27].

Personality prediction is an important phase of personality-aware recommender systems, and the two main methods for doing so is through questionnaires and automated means. Generally, questionnaires are more accurate in assessing personality; however, the process is tedious while the automated approach is easier to conduct, by utilising user's existing data that can be either text, images, videos, likes (behavioural data) etc. [8] Predicting personality from text is a popular automated approach that is based on personality theory claiming that words can reveal some psychological states and personality of the author of the text. There are two main categories of techniques, the feature-based and the deep learning: the former uses unigrams/n-grams (open vocabulary approach) or lexicons (closed vocabulary) of features relevant to personality, and the latter text embeddings learned from large corpus of text in an unsupervised manner (language models). Popular feature-based methods utilize the Mairesse [20] and linguistic inquiry and word count techniques [28]. Features from these are fed into different machine learning classifiers (e.g., Naïve Bayes, support vector machines) to make predictions. Obtaining such features however is a costly process and cannot effectively represent the original text semantics. To avoid feature engineering, deep neural models and language models are employed to learn text representations that currently result in improved accuracy. Deep models focus on the context of the text and not just a static representation for a word or a sentence. Those kind of deep learning techniques are using an attention mechanism that focuses on giving weights to words based on how they are used in a text giving the ability of capturing the semantic content [15]. A popular architecture is the BERT (Bidirectional Encoder Representations from Transformers) that utilizes transformers neural network architecture. Attention-based transformers have shown that collecting the semantics of a text improves the performance level and the predication accuracy of ML personality models [13]. Given this, the method proposed in this paper utilizes attention-based personality prediction.

Most approaches use a binary classifier for each of the personality traits (MBTI) such as a classifier for extraversion-introversion etc. Such methods require pre labeled data with the personality class. The first step in the process is the vectorization of the text into a form that can be processed by ML algorithms [34]. This can be done using open/closed lexicons or sentence embeddings in the case of deep learning methods (BERT). The vectorized data is used to train a classifier using the data label or fine tune a pretrained model in the case of BERT. The trained and validated model can be used to predict unseen data. Recent personality classification techniques that utilize deep learning for Big Five personality prediction, such as the DeepPerson [40] demonstrate classification performance (AUC score) of around 70% per personality dimension, using different training datasets, which is much lower compared to classifiers that use MBTI data.

3 TECHNICAL BACKGROUND

The proposed method utilises a named entity recognition to extract food preferences, an automated eWOM topic modelling for the identification of themes discussed in review's text, a BERT-based personality classification, and an ensemble tree-based regression for the prediction of consumer restaurant ratings.

3.1 User's Food Preference Extraction

A named entity recognition (NER) is utilized to extract the food preferences of customers. NER is a major component in NLP systems to extract information from unstructured text. An entity can be any word or sentence that refers to the concept of question. There are two main approaches for the creation of a NER, the model-based, and rule-based approach. The latter focuses on the grammatical rules and linguistic terms to extract entities. The model-based approach generates machine learning models using a text with prelabelled entities. Most food NER models as reported in [29] are trained on data that did not include Cyprus dishes thus their predictions were insufficient for our case. There is a wide range of generic libraries suitable for NER, such as NLTK, spaCy, Stanford NER, Stanza, and Flair, but none was able to provide appropriate food recognition in text.

To extract food preferences, the spacy library was utilized, and several rules have been specified that enabled the extraction of sentences that mention food consumption such as "I ate ","I had for dinner" etc. To generate a sufficient training set, a local and international food recipe dataset was used. The returned sentences were annotated automatically based on the position in the sentence where the food entity occurred. This was necessary in order to create a training dataset labelled with food names, and their start/end position in the sentence, based on which the food NER training was performed. The trained NER achieved an overall accuracy of 81% (70/30 train-test split) and was applied on the restaurant reviews to extract foods associated with each review. Due to the large number of food entities that were generated, there were a lot of repetitions due to different spellings. Thus, to reduce the dimensionality of the dataset, a feature selection process was performed using a random forest machine learning model to identify the most important food names using the review ratings as the target variable. The process yielded the optimum number of features (220) that resulted in the best model performance. The selected food features where then one-hot encoded for each consumer review. To identify the food preferences of each user, reviews were grouped by user and the most frequent food entities in each user's reviews were considered as food preferences. This process considers that, when customers visit different restaurants and write comments about the food they ordered, irrespective of the food's quality and the review rating, it constitutes food preference of the user.

3.2 EWOM Topic Modelling

Topic modelling is a popular tool for extracting information from unstructured data and is used in this work to identify themes discussed by consumers in eWOM. Topic models generally involve a statistical model aiming at finding topics that occur in a collection of documents [25]. Two of the most popular techniques for topic analysis are the Latent Dirichlet Allocation and the Structural Topic Model (STM). In this study, the STM approach [30] is used to develop a topic model due to its ability to incorporate reviews' metadata such as sentiment (rating>3) that help with interpreting and naming the identified topics. Each topic in STM represents a set of words that occur frequently together in a corpus and each document is associated with a probability distribution of topics per document. The process for learning the topic model initiates with data preprocessing that includes removal of common and custom stop-words and irrelevant information (punctuation), followed by tokenization (breaking sentences into word tokens), and stemming (converting words to their root form). Initially, common stop-words were considered and gradually with the refinement of the model, additional stop-words that were irrelevant to our goal were added to the list of custom stop-words such as names of people, restaurants, cities, etc. The optimum number of topics that best fits the dataset is identified through an iterative process examining different values for the number of topics (K) and inspecting the semantic coherence, held out likelihood, and exclusivity of the model at each iteration until a satisfactory model is produced [30]. Coherence measures the degree of semantic similarity between high scoring words in the topic. Held out likelihood tests a trained topic model using a test set that contains previously unseen documents. Exclusivity measures the extent to which top words in one topic are not top words in other topics. The naming of the topics was performed manually based on domain knowledge and the most prevalent words that characterize each topic.

3.3 BERT Personality Classification

Recent benefits of the "attention" mechanism in deep learning models have demonstrated state-of-the-art performance in numerous text analysis tasks such as classification.

BERT uses a multi-layer bidirectional transformer encoder and is inspired by the concept of knowledge transfer, since in many problems it is difficult to access sufficiently large volume of labelled data to train deep models. In transfer learning, a pre-train a model is learned from massive unlabeled datasets not representing the target problem, but allows the learning of general knowledge. BERT-like approaches provide pretrained models and their embedded knowledge can be transferred to a target domain where labelled data is limited. Fine-tuning such models is performed using a labelled dataset representing the actual problem; these tune the model to the task at hand. Fine-tuning adds a feedforward layer on top of the pre-trained BERT. Previous work has demonstrated that this pre-training and fine-tuning approach outperforms existing text classification approaches. In our case, fine-tuning the BERT model was performed using publicly available personality labelled data. BERT has been used for personality prediction using the Personality Cafe MBTI dataset in [17] achieving an accuracy of around 0.75. In contrast, other deep learning methods that use the Big 5 model as well as the popular stream-of-consciousness essay dataset such as the one reported in [21] using CNN, achieve inferior classification performance.

Despite their good results, BERT-based approaches have been criticized that their best performance is reported with short texts. Long text refer to text with more than 512 tokens. Such text however are computationally expensive to process thus most transformers models limit the number of tokens they can process simultaneously. In our case, most reviews produced by consumers exceeded the 512 tokens limit and thus the prediction of personality was considered as a long text classification problem. Different methods exist to dealing with this issue, which include the naïve head-only, tail only or semi-naïve approaches, that either use the top number of words, bottom number of words, or combination of top/bottom/important words in the text. Such approaches lose information but have a minimum computational cost. Recent works have sought to alleviate the computational cost constraint by applying more sophisticated models to longer text instances such as dividing the long text into chunks and combining the embeddings of the chunks. However, work by Sun et al. [36] that investigated different long-text treatment methods for consumer reviews, showed that the best classification performance is achieved using naïve methods such as using only the head or tail tokens of the text while dropping all other content. In this work, we explore the naïve and semi naïve methods to find the one with the best personality classification performance prior to labelling users with their personality. The results, described in a subsequent section, show that the naïve approach yielded the best performance, which is in line with [36].

3.4 XGBoost Regression

XGBoost regression is used in this study due to its ability of producing good results in similar problems. It is an ensemble method; hence multiple trees are constructed with the training of each tree depending on errors from previous trees' predictions. Gradient

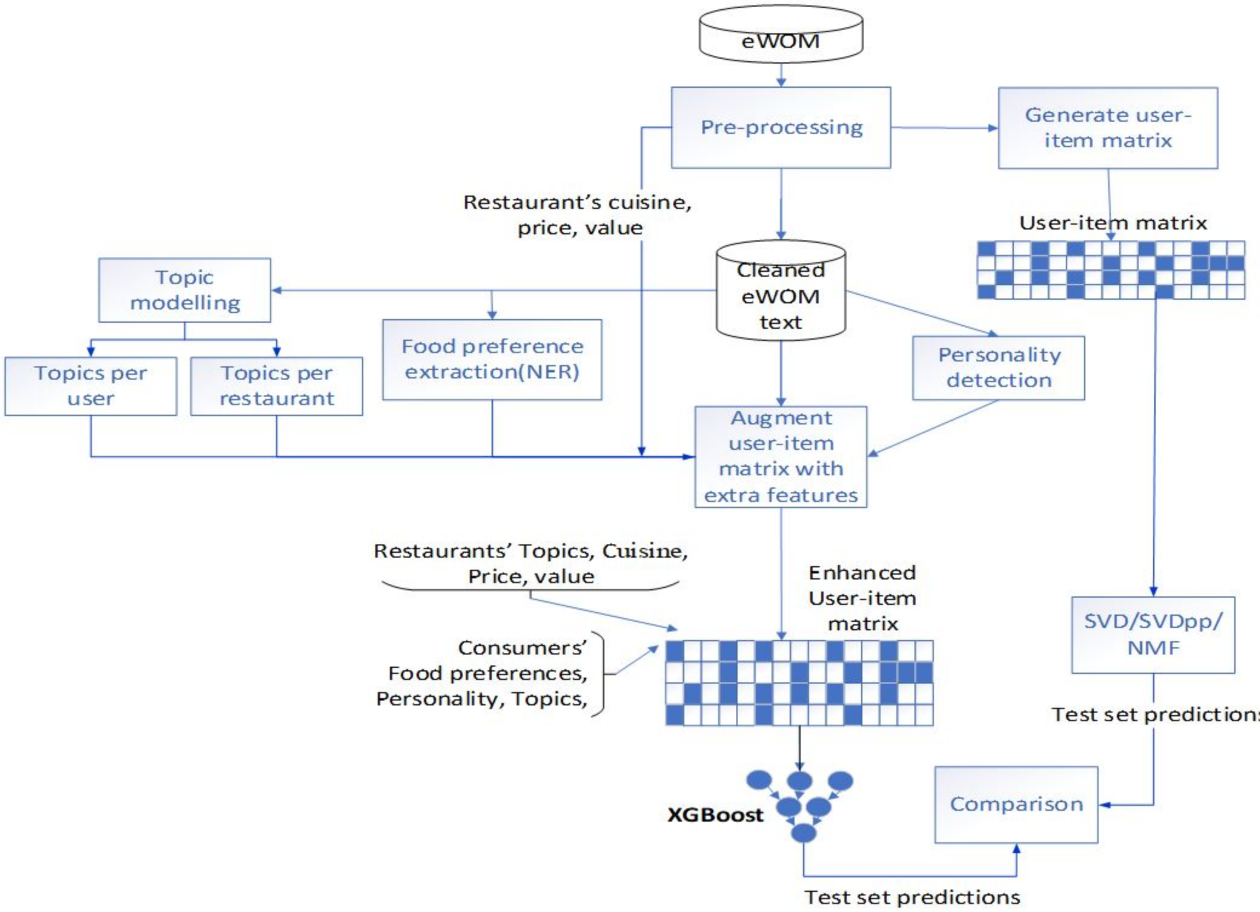

Figure 1: Overview of the approach and its evaluation

descent is used to generate new trees based on all previous trees while optimizing for loss and regularization. XGBoost regularization component balances complexity of the learned model against predictability. XGBoost optimization is required to minimize model overfitting and treating data imbalance, by tuning multiple hyperparameters. The optimal values of hyperparameters can be determined with different techniques such as the exhaustive (grid search), Bayesian, or random. The grid search method combines all possible values of each parameter, to obtain the model with best performance, while the Bayesian utilizes results from previous optimization cycles to identify hyperparameters values with higher probability in improving the classifiers performance. Grid search is better but slower while Bayesian is faster but not as accurate. In this work, the grid search approach is adopted.

4 METHODOLOGY

The methodology employed to address our research question is presented in Figure 1 and is implemented via the following steps.

(1) Collection of restaurant reviews from TripAdvisor and extraction of consumers' eWOM and additional explicit information of restaurants such as cuisine type, price range, and value for money;
(2) Preprocessing of the data and preparation for subsequent analyses (topic modelling, personality classification). Preprocessing includes punctuations and URLs elimination, lowering of text, stop words removal, tokenization, stemming, and lemmatization. During this step, the user-item matrix is generated with rows corresponding to consumers and columns to restaurants. The cells of the matrix contain ratings when these are available since customers did not visit all restaurants;
(3) Development of a topic model using as corpus the eWOM (reviews) to identify consumers' opinions and how these are associated with each review. Restaurant's topics are generated by averaging the topics theta values associated with each restaurant. This represents common consumer opinions per restaurant;
(4) Assessment of customers' personality from eWOM is achieved using the MBTI BERT personality classification model;

(5) Food preferences of users are extracted from eWOM's text using a custom NER model;
(6) The explicit information from each restaurant is combined with implicit information that emerges from personality analysis, food preference, and topic modelling. These features are used collectively to enhance the user-item matrix and are used to train an Extreme Gradient Boosting (XGboost) regressor model using as output variable the user rating of restaurants and taking values in the range [1-5]. The XGBoost is optimized using hyperparameter tuning and validated using train/test data split (70/30). The trained model is used to predict user ratings for restaurant users have never visited;
(7) The performance of the XGBoost model is compared against that of three popular model-based CF techniques, namely SVD, SVD++, and NMF. The comparison models are trained using the initial user-item matrix while the XGBoost using the enhanced user-item matrix that includes explicit and implicit information. The performance of the models is assessed using popular evaluation metrics such as mean absolute error (MAE), mean squared error (MSE), and root mean squared error (RMSE).

5 RESULTS

The data collected includes 105k reviews written in English from tourists who visited Cyprus between 2010 to 2020 and posted reviews about their experience with restaurants in Cyprus (publicly available). The total number of unique users were 56800 and the number of restaurants were 650. Figure 2 depicts descriptive statistics of reviews ratings per year. For this study only users with at least 20 reviews are considered and only restaurants with at least 50 reviews yielding 93 unique users and 410 venues.

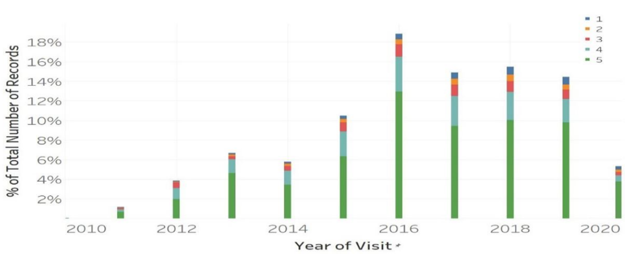

Figure 2: Percentage of restaurant review ratings [1-5] per year from 2010 to 2020

5.1 Learned Topic Model

To extract consumers' discussed themes from eWOM, an STM topic model was developed using the estimated optimum K (30) number of topics based on the model's performance metrics in Figure 3, with focus on high coherence, high held-out likelihood, low residuals, and high lower bound scores.

The naming of the topics in Table 1 was based on domain knowledge, words with highest probability in each topic and words with high Lift score. Lift gives higher weight to words that appear less frequently in other topics.

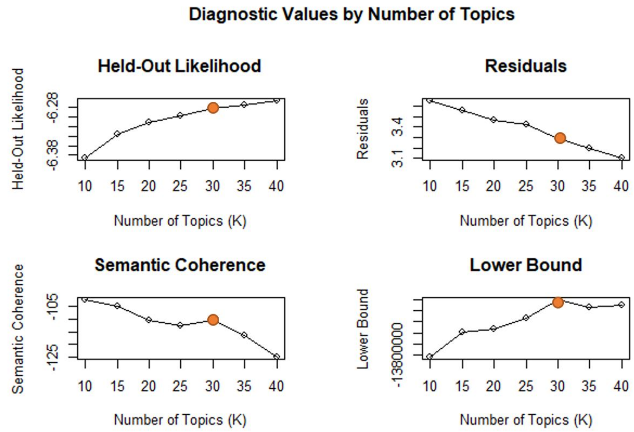

Figure 3: Topic performance measures for identifying the optimum number of topics. The red circle indicated the K number of topics selected

The probability distribution of topics per review denotes the probability of each topic discussed in a review and the sum of all topics' probabilities in each review total 1. Reviews are associated with the distribution of topics prevalence per review. The trained STM model's theta values per review refer to the probability that a topic is associated with each review. These theta values, shown in Figure 4, were used as features during the training of the XGBoost model along with other features.

5.2 Personality Labelling

The training of the binary classifiers was performed using the Personality Cafe MBTI dataset consisting of joint user posts on a social network labeled by personality type defined using MBTI questionnaire. The dataset is publicly available on "Kaggle" [14]. To identify the BERT long text approach with the best classification performance, two techniques were examined, namely the naïve and semi naïve approach and the one with the best performance was used in the workflow. For the naïve approach, we used the head-only using as sentence length the 256 and 512 words and for the semi-naïve we used chunking of text into 128 words and combining their embedding. The results from this process, presented in Table 2, showed that the 512-naïve-head approach outperformed the other approaches and thus it was employed in users' personality classification. Results from the BERT model outperformed personality models trained using the same dataset and thus improved our confidence in the personality prediction of each user.

The personality distributions in Figure 5 show descriptive statistics regarding the personalities of users according to the detected personality from the MBTI BERT classifier fine-tuned using a labelled datasets and treating long text using the naïve-head approach with 512 tokens. The acronyms refer to combination of dimensions of the MBTI model. The trained BERT model predicts for each dimension of the personality model the probability that a user belongs to any of the personality traits (i.e., probability for Extraversion

Table 1: Specified names for the topics that emerged from STM analysis

Topics	Words with high probability and lift scores	Topic Name
Topic1:	great, really, music, live, day, atmosphere, enchiladas, music, really, live, pub	Entertainment Atmosphere
Topic2:	nice, prices, atmosphere, reasonable, big, family, polite, quick, nice, relaxing, families, cafe	Family Restaurant
Topic3:	time, excellent, went, night, amazing, first, every, occasions, stay, went, amaze, week	Special Occasion
Topic4:	eat, new, end, found, places, second, thai, always, second, none	New Place
Topic5:	lovely, recommend, highly, enjoyed, beautiful, setting, party, setting, party, hosts, fabulous, thoroughly, absolutely	Party Place
Topic6:	well, lunch, local, attentive, wonderful, presented, chose, breaks, attentive, presented, chose	Lunch
Topic7:	evening, bar, friends, though, group, customers, quiet, whiskies, though	Evening/Bar
Topic8:	restaurant, location, must, beach, view, right, perfect, definitely, must, far	Location
Topic9:	visit, will, back, really, worth, definitely, going, visit, called	Worth Visiting
Topic10:	wedding, amazing, even, similar, impression, organize, guest, events, beyond, pleasure	Wedding Place
Topic11:	many, birthday, soon, booked, kitchen, also, october, good, love see, year this, flight, travel, celebration	Celebration parties
Topic12:	experience, nothing, special, whole, maybe, dining, perfection, fiancée, maybe	Not Worth
Topic13:	restaurant, probably, also, mountains, open, available, well, best, more, owner, managers, troodos	Out of town
Topic14:	summer, use, even, range, late, cool, evenings, use, dine, cozy	Summer location
Topic15:	always, can, class, owners, restaurant, number, first, classy, number, varied, feeling, interesting, hidden	Fabulous Place
Topic16:	two, outside, can, inside, sit, world, get, disappointed, aircon, magic, noise, traffic, heat	Outside eating
Topic17:	thai, tourist, across, gem, trying, partner, avoid, duck, again, overall, always, bespoke, gimmicks, hardcore	Asian Cuisine
Topic18:	bit, little, better, average, like, quite, expensive, much, however, criticisms, average,	Average Place
Topic19:	different, small, cheese, also, breakfast, euros, greek, platter, platter, options, vegetarian, bacon, eggs	Breakfast
Topic20:	wife, return, disappointed, restaurant, reviews, favourite, holiday, isn't, trip, done	Disappointment
Topic21:	old, cypriot, road, stop, village, along, street, waitresses, road, walk	Stop during trips
Topic22:	busy, get, people, table, lot, need, without, joyful, early, book	Busy Place
Topic23:	years, restaurant, made, visiting, coming, several, since, ago, forward, visits	Visit over years
Topic24:	staff, friendly, always, see, come, welcoming, feel, chat, truly, smile, come	Welcoming Staff
Topic25:	chips, served, priced, set, large, course, portion, adults, chips, portion, reasonably	Good portions
Topic26:	best, cyprus, don't, ever, never, restaurants, know, traditional, meze	Traditional foods
Topic27:	value, money, recommended, excellent, variety, high, meals , bringing, best, cyprus, eaten	Value for money
Topic28:	couldn't, enough, friend, eat, fresh, away, wow, take, out, basilica, excellent, lovely, rice, more, time ,highly	Fresh ingredients
Topic29:	ordered, came, table, order, waiter, asked, arrived, minutes, waited, waitress, left, seated, orders, bill	Bad service
Topic30:	just, restaurant, basic, way, like, standard, much, full, unacceptable, much, restaurant	Nothing Special

Table 2: Performance results per long text treatment

Long text treatment for MBTI BERT	AUC	ACC
Naïve - head 512 tokens	0.878	0.839
Naïve - head 256 tokens	0.784	0.759
Semi naïve - Sliced text 128 tokens	0.653	0.662

– Introversion (E/I), Sensing – Intuition (S/N), Thinking – Feeling (T/F), and Judging – Perceiving (J/P)). Combinations of letter from each category generate 16 four-letter personality types: ISFJ, INFP, INFJ, ISTP, ISTJ, ISFP, INTP, INTJ, ENTP, ESFP, ENFP, ESFJ, ESTP, ESTJ, ENFJ and ENTJ, the distribution of which is depicted in Figure 5. The BERT model deals with 4 classifiers, one for each of the dimensions above. The classifier's average area under the curve (AUC) performance is 87%. This is an improved performance compared to alternative personality classification techniques that utilize deep learning and Big Five personality model, such as the DeepPerson [40] that achieved AUC of around 70%.

5.3 Training and Evaluating the XGBoost Models

The enhanced user-item matrix that emerged from the personality model, food preferences, and the topics associations per user and venue were used to train an XGB regressor model.

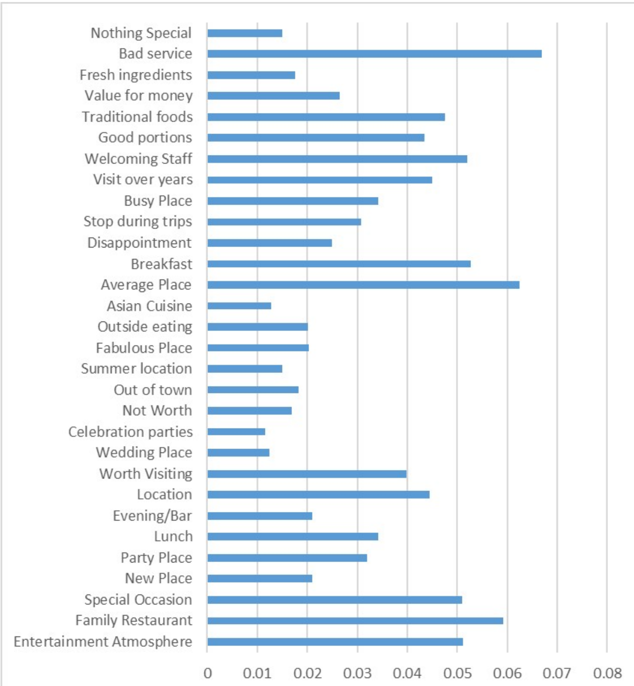

Figure 4: Average theta values per topic

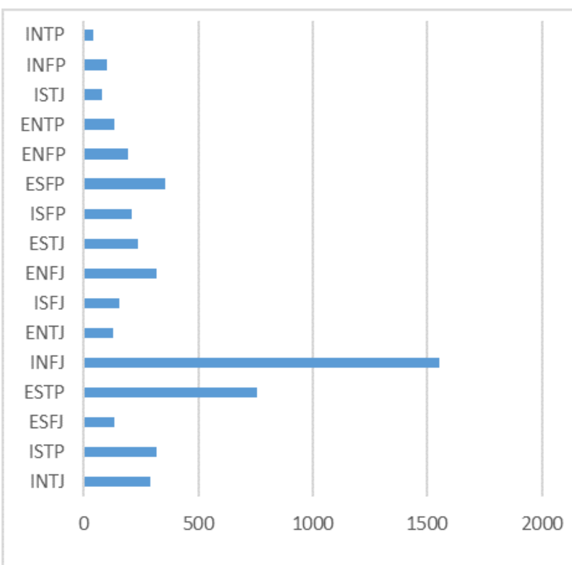

Figure 5: Distribution of MBTI personality traits using each dimension's acronyms

The XGBoost model underwent hyperparameter tuning prior to training by tuning the models' learning rate, gamma, subsample, and regularization options using grid search. Traditional recommendation models, namely SVD, SVD++, and NMF were generated using the surprise python library. The models were compared based on the following performance metrics: the mean absolute error (MAE) that represents the average of the absolute difference between the real and predicted values, Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) that is the square root of MSE. Comparison of the two models against traditional recommendation techniques revealed an improved performance of the personality-based approaches over these baseline models. The traditional techniques were also optimized by tuning two hyperparameters, the number of factors and the regularization value.

In the experiments conducted using the aforementioned restaurants reviews, the data was initially split into test and training sets (70/30) using stratified sampling to guarantee that all user ratings are sufficiently represented in the test and training samples. The models were hyper tuned, trained, and tested using the same samples. The aforementioned metrics were computed, and the results that emerged (see Table 3) show that the MBTI XGBoost model produced the best performance among all other models. Both personality-based models outperformed traditional approaches, which indicates that the use of personality and eWOM-extracted topics improved the recommendations.

Table 3: Performance results per model incorporating all features

Performance metric (lower is better)	**SVD**	**SVD++**	**NFM**	**XGB**
Mean Absolute Error (MAE)	0.65	0.68	0.82	0.40
Mean Squared Error (MSE)	0.87	0.89	1.22	0.24
Root Mean Squared Error (RMSE)	0.93	0.94	1.10	0.49

6 CONCLUSIONS

This study proposes a combined user-preference with user personality restaurant recommendation approach and constitutes one of the first studies that use customer preference along with personality in the restaurant recommendation problem. It utilizes a popular personality model (MBTI) to enhance the restaurant recommendation process by fine tuning a BERT classification model on personality labelled dataset. Due to the length of the training data, the best long-text handling approach (naïve-head 512) was employed during BERT model tuning. EWOM themes are extracted through topic modelling from eWOM's text and are also used as additional features of restaurants and users that refer to implicit preferences of users and properties of restaurants. All aforementioned features are used collectively to train an XGBoost regressor to predict consumers' satisfaction (i.e., rating) for unvisited restaurants. The results show that the MBTI model in combination with topics from eWOM outperforms the model-based collaborative filtering techniques, offering a first indication that the application of personality and food preferences in restaurant recommendation can have valuable results. Future work will focus on evaluating additional long-text handling techniques and combine the results of the learned classifiers with other traditional machine learning models in an ensemble manner to improve further the performance of personality classification, given that personality is a valuable feature that enhances restaurant recommendation.

REFERENCES

[1] Mohammad Hossein Amirhosseini and Hassan Kazemian. 2020. Machine learning approach to personality type prediction based on the Myers–Briggs type indicator®. *Multimodal Technologies and Interaction* 4, 1 (2020), 9. https://doi.org/10.3390/mti4010009

[2] Carl Anderson. 2018. A survey of food recommenders. *CoRR* abs/1809.0 (2018). http://arxiv.org/abs/1809.02862

[3] Asim Ansari, Skander Essegaier, and Rajeev Kohli. 2000. Internet Recommendation Systems. *Journal of Marketing Research* 37, 3 (aug 2000), 363–375. https://doi.org/10.1509/jmkr.37.3.363.18779

[4] Elham Asani, Hamed Vahdat-Nejad, and Javad Sadri. 2021. Restaurant recommender system based on sentiment analysis. *Machine Learning with Applications* 6 (2021), 100114. https://doi.org/10.1016/j.mlwa.2021.100114

[5] Gregory J Boyle. 1995. Myers-Briggs Type Indicator (MBTI): Some Psychometric Limitations. *Australian Psychologist* 30, 1 (1995), 71–74. https://doi.org/10.1111/j.1742-9544.1995.tb01750.x

[6] Fabio Celli and Bruno Lepri. 2018. Is big five better than MBTI? A personality computing challenge using Twitter data. *CEUR Workshop Proceedings* 2253 (2018).

[7] Evripides Christodoulou, Andreas Gregoriades, Maria Pampaka, and Herodotos Herodotou. 2022. Personality-Informed Restaurant Recommendation. In *World Conference on Information Systems and Technologies*, Alvaro Rocha, Hojjat Adeli, Gintautas Dzemyda, and Fernando Moreira (Eds.). Springer International Publishing, Cham, 13–21.

[8] Sahraoui Dhelim, Nyothiri Aung, Mohammed Amine Bouras, Huansheng Ning, and Erik Cambria. 2022. A Survey on Personality-Aware Recommendation Systems. *Artif. Intell. Rev.* 55, 3 (mar 2022), 2409–2454. https://doi.org/10.1007/s10462-021-10063-7

[9] John Gountas and Sandra Gountas. 2007. Personality orientations, emotional states, customer satisfaction, and intention to repurchase. *Journal of Business Research* 60, 1 (2007), 72–75. https://doi.org/10.1016/j.jbusres.2006.08.007

[10] Andreas Gregoriades, Maria Pampaka, and Michael Georgiades. 2020. A Holistic Approach to Requirements Elicitation for Mobile Tourist Recommendation Systems. In *Future of Information and Communication Conference*, Kohei Arai and Rahul Bhatia (Eds.). Springer International Publishing, Cham, 857–873.

[11] Sindhu B Hegde, Supriya Satyappanavar, and Shankar Setty. 2018. Sentiment based Food Classification for Restaurant Business. In *2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*. 1455–1462. https://doi.org/10.1109/ICACCI.2018.8554794

[12] Guntur Budi Herwanto and Annisa Maulida Ningtyas. 2017. Recommendation system for web article based on association rules and topic modelling. *Bulletin of Social Informatics Theory and Application* 1, 1 (2017), 26–33. https://doi.org/10.31763/businta.v1i1.36

[13] He Jun, Liu Peng, Jiang Changhui, Liu Pengzheng, Wu Shenke, and Zhong Kejia. 2021. Personality Classification Based on Bert Model. *Proceedings of 2021 IEEE International Conference on Emergency Science and Information Technology, ICESIT 2021* (2021), 150–152. https://doi.org/10.1109/ICESIT53460.2021.9697048

[14] Kaggle. 2017. (MBTI) Myers-Briggs Personality Type Dataset. https://www.kaggle.com/datasets/datasnaek/mbti-type

[15] Spyridon Kardakis, Isidoros Perikos, Foteini Grivokostopoulou, and Ioannis Hatzilygeroudis. 2021. Examining attention mechanisms in deep learning models for sentiment analysis. *Applied Sciences (Switzerland)* 11, 9 (2021). https://doi.org/10.3390/app11093883

[16] Raghav Pavan Karumur, Tien T. Nguyen, and Joseph A. Konstan. 2018. Personality, User Preferences and Behavior in Recommender systems. *Information Systems Frontiers* 20, 6 (dec 2018), 1241–1265. https://doi.org/10.1007/s10796-017-9800-0

[17] Sedrick Scott Keh and I-Tsun Cheng. 2019. Myers-Briggs Personality Classification and Personality-Specific Language Generation Using Pre-trained Language Models. https://doi.org/10.48550/ARXIV.1907.06333

[18] Yeong Gug Kim, Anita Eves, and Caroline Scarles. 2009. Building a model of local food consumption on trips and holidays: A grounded theory approach. *International Journal of Hospitality Management* 28, 3 (2009), 423–431. https://doi.org/10.1016/j.ijhm.2008.11.005

[19] Yehuda Koren, Robert Bell, and Chris Volinsky. 2009. Matrix Factorization Techniques for Recommender Systems. *Computer* 42, 8 (2009), 30–37. https://doi.org/10.1109/MC.2009.263

[20] François Mairesse, Marilyn A. Walker, Matthias R. Mehl, and Roger K. Moore. 2007. Using linguistic cues for the automatic recognition of personality in conversation and text. *Journal of Artificial Intelligence Research* 30 (2007), 457–500. https://doi.org/10.1613/jair.2349

[21] Navonil Majumder, Soujanya Poria, Alexander Gelbukh, and Erik Cambria. 2017. Deep Learning-Based Document Modeling for Personality Detection from Text. *IEEE Intelligent Systems* 32, 2 (2017), 74–79. https://doi.org/10.1109/MIS.2017.23

[22] Sushma Malik, Anamika Rana, and Mamta Bansal. 2020. A Survey of Recommendation Systems. *Information Resources Management Journal* 33, 4 (oct 2020), 53–73. https://doi.org/10.4018/IRMJ.2020100104

[23] Kye-Hong Min and Timothy Jeonglyeol Lee. 2014. Customer Satisfaction with Korean Restaurants in Australia and Their Role as Ambassadors for Tourism Marketing. *Journal of Travel & Tourism Marketing* 31, 4 (2014), 493–506. https://doi.org/10.1080/10548408.2013.877412

[24] Senthilselvan Natarajan, Subramaniyaswamy Vairavasundaram, Sivaramakrishnan Natarajan, and Amir H. Gandomi. 2020. Resolving data sparsity and cold start problem in collaborative filtering recommender system using Linked Open Data. *Expert Systems with Applications* 149 (2020). https://doi.org/10.1016/j.eswa.2020.113248

[25] Sergey I. Nikolenko, Sergei Koltcov, and Olessia Koltsova. 2017. Topic modelling for qualitative studies. *Journal of Information Science* 43, 1 (feb 2017), 88–102. https://doi.org/10.1177/0165551515617393

[26] Mehrbakhsh Nilashi, Othman bin Ibrahim, Norafida Ithnin, and Nor Haniza Sarmin. 2015. A multi-criteria collaborative filtering recommender system for the tourism domain using Expectation Maximization (EM) and PCA–ANFIS. *Electronic Commerce Research and Applications* 14, 6 (oct 2015), 542–562. https://doi.org/10.1016/j.elerap.2015.08.004

[27] Sampo V. Paunonen. 2003. Big Five Factors of Personality and Replicated Predictions of Behavior. *Journal of Personality and Social Psychology* 84, 2 (2003), 411–424. https://doi.org/10.1037/0022-3514.84.2.411

[28] James W Pennebaker, Martha E Francis, and Roger J Booth. 2001. Linguistic inquiry and word count: LIWC 2001. *Mahway: Lawrence Erlbaum Associates* 71, 2001 (2001), 2001.

[29] Gorjan Popovski, Barbara Koroušić Seljak, and Tome Eftimov. 2020. A Survey of Named-Entity Recognition Methods for Food Information Extraction. *IEEE Access* 8 (2020), 31586–31594. https://doi.org/10.1109/ACCESS.2020.2973502

[30] Margaret E. Roberts, Brandon M. Stewart, Dustin Tingley, Christopher Lucas, Jetson Leder-Luis, Shana Kushner Gadarian, Bethany Albertson, and David G. Rand. 2014. Structural topic models for open-ended survey responses. *American Journal of Political Science* (2014). https://doi.org/10.1111/ajps.12103

[31] Jesús F. Salgado. 2002. The Big Five personality dimensions and counterproductive behaviors. *International Journal of Selection and Assessment* 10, 1-2 (2002), 117–125. https://doi.org/10.1111/1468-2389.00198

[32] Zeinab Shahbazi, Yungcheol Byun, and Yung-Cheol Byun. 2020. Product Recommendation Based on Content-based Filtering Using XGBoost Classifier. *International Journal of Advanced Science and Technology* 29, 04 (2020), 6979–6988. https://www.researchgate.net/publication/342864588

[33] Nícollas Silva, Diego Carvalho, Adriano C.M. Pereira, Fernando Mourão, and Leonardo Rocha. 2019. The Pure Cold-Start Problem: A deep study about how to conquer first-time users in recommendations domains. *Information Systems* 80 (2019), 1–12. https://doi.org/10.1016/j.is.2018.09.001

[34] Sanja Stajner and Seren Yenikent. 2020. A Survey of Automatic Personality Detection from Texts. In *Proceedings of the 28th International Conference on Computational Linguistics*. 6284–6295. https://doi.org/10.18653/v1/2020.coling-main.553

[35] Wen-Zhong Su and Po-Hsien Lin. 2016. A Study of Relationship Between Personality and Product Identity. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Vol. 9741. 266–274. https://doi.org/10.1007/978-3-319-40093-8_27

[36] Chi Sun, Xipeng Qiu, Yige Xu, and Xuanjing Huang. 2019. How to Fine-Tune BERT for Text Classification? https://doi.org/10.48550/ARXIV.1905.05583

[37] Lihua Sun, Junpeng Guo, and Yanlin Zhu. 2019. Applying uncertainty theory into the restaurant recommender system based on sentiment analysis of online Chinese reviews. *World Wide Web* 22, 1 (2019), 83–100. https://doi.org/10.1007/s11280-018-0533-x

[38] Hanfei Wang, Yuan Zuo, Hong Li, and Junjie Wu. 2021. Cross-domain recommendation with user personality. *Knowledge-Based Systems* 213 (2021), 106664. https://doi.org/10.1016/j.knosys.2020.106664

[39] Wen Wu, Li Chen, and Yu Zhao. 2018. Personalizing recommendation diversity based on user personality. *User Modeling and User-Adapted Interaction* 28, 3 (aug 2018), 237–276. https://doi.org/10.1007/s11257-018-9205-x

[40] Kai Yang, Raymond YK Lau, and Ahmed Abbasi. 2022. Getting Personal: A Deep Learning Artifact for Text-Based Measurement of Personality. *Information Systems Research* (2022).

[41] Chenbin Zhang, Hongyu Zhang, and Jianqiang Wang. 2018. Personalized restaurant recommendation method combining group correlations and customer preferences. *Information Sciences* 454-455 (jul 2018), 128–143. https://doi.org/10.1016/j.ins.2018.04.061