



Predicting users' review helpfulness: the role of significant review and reviewer characteristics

Muhammad Shahid Iqbal Malik¹

Published online: 19 February 2020
© Springer-Verlag GmbH Germany, part of Springer Nature 2020

Abstract

There are thousands of reviews constantly being posted for popular products on e-commerce websites. The number of reviews is rapidly increasing which creates information overload problem. To solve this problem, many websites introduced a feedback mechanism to vote for a review (helpful or not). The attracted votes reflect the review helpfulness. This study addresses the review helpfulness prediction problem and investigated the impact of review, reviewer and product features. Multiple helpfulness prediction models are built using multivariate adaptive regression, classification and regression tree, random forest, neural network and deep neural network approaches using two real-life Amazon product review datasets. Deep neural network-based review helpfulness prediction model has outperformed. The results demonstrate that review-type characteristics are most effective indicators as compared to reviewer and product type. In addition, hybrid combination (review, reviewer and product) of proposed features demonstrates the best performance. The influence of product type (search and experience) on review helpfulness is also examined, and reviews of search goods show strong relationship to review helpfulness. Our findings suggest that polarity of review title, sentiment and polarity of review text and cosine similarity between review text and product title effectively contribute to the helpfulness of users' reviews. Reviewer production time and reviewer active since features are also strong predictors of review helpfulness. Our findings will enable consumers to write useful reviews that will help retailers to manage their websites intelligently by assisting online users in making purchase decisions.

Keywords Deep neural network · Cosine similarity · Product reviews · Reviewer production time · Review helpfulness

1 Introduction

Online customer reviews (OCRs) which refer to the customer-generated feedback emerged as electronic word of mouth (eWOM) for online users of current generation and e-commerce retailers (Bertola and Patti 2016). Nowadays, customer reviews are assumed to be a major source of opinions and evaluations of online product, and therefore in e-commerce, both practitioners and academics have paid more attention to understand the role of customer reviews (Duan et al. 2008; Forman et al. 2008; Samha and Li 2014). Chen and Xie (2008) reveal that customer reviews are an

important element of marketing communication mix and strongly affect the quality improvement in business organizations and product sales. Online users usually preferred other users' evaluations and feedbacks before making purchase decisions (Zhang and Piramuthu 2016). Recently, web retailers are using customer reviews to understand the attitudes and opinions of online users. Therefore, customer reviews are being given more attentions and significance by the business retailers because they are a thread or an opportunity for e-commerce businesses (Li and Hitt 2010; Anderson and Magruder 2012; Yan et al. 2015). However, the amount of customer reviews is rapidly increasing for popular products that result in information overload problem (Liu and Huang 2008). Such a large amount of reviews are being considered as a big data challenge for customers and online retailers (Xu and Xia 2008; Samha and Li 2014).

To solve this problem, many websites introduced a feedback mechanism to vote for a review as helpful or not.

Communicated by V. Loia.

✉ Muhammad Shahid Iqbal Malik
shahid.iqbal@cust.edu.pk; msi_id@yahoo.com

¹ Department of Computer Science, Capital University of Science and Technology, Islamabad, Pakistan

The attracted votes reflect the helpfulness of online reviews. Alternatively, helpfulness is the ratio of the helpful votes to the total votes received by a review. It is the target variable in this research. Early studies used the basic indicators to predict the helpfulness of user reviews (Pang and Lee 2002; Otterbacher 2009). The indicators are review length, review star rating and length, thumbs up/down. Later, few studies focused to investigate the qualitative measures in addition to the quantitative ones such as reviewer experience, reviewer impact, linguistics features and reviewer cumulative helpfulness to analyze review helpfulness (Mudambi and Schuff 2010; Huang et al. 2015; Krishnamoorthy 2015; Qazi et al. 2016). But, majority of prior studies utilized reviewer, review and product-type characteristics in the limited scope. There exist a number of significant indicators of review and reviewer types, which are not part of state-of-the-art approaches.

2 Motivation and research contributions

An effective machine learning (ML) model is always needed to investigate the significance of various types of indicators on online review helpfulness. Deep learning is one of the broader families of ML models which are based on learning representation of data. Numerous architectures of deep learning such as recurrent neural networks, deep belief networks, deep neural networks and convolutional deep neural networks had already been utilized in multiple domains and demonstrated promising results for various problems. The deep neural network is a feedforward neural network that has multiple hidden layers between the input and output layers. It can model complex nonlinear relationships and are able to resolve new problems relatively easily. Since deep neural network performed quite effective in diverse domains, therefore, it is being adapted as a model of choice for desirable predictive performance.

Due to rapid increase in the collection of customer reviews on e-commerce platforms, different factors have varied effect on helpfulness of online reviews. Sentiments and emotions embedded in the review text have varied effects on review helpfulness (Chua and Banerjee 2016; Malik and Hussain 2017). With the passage of time, new challenges and opportunities are being encountered. There is a need to address the new dimensions of textual and reviewer indicators of product reviews. Prior studies explored the influence of review sentiment on helpfulness (Chua and Banerjee 2016), product recommendation (Zhang and Piramuthu 2016) and order effect on review helpfulness (Zhou and Guo 2017). This research addresses the following research questions:

- Which machine learning algorithm delivers the best predictive performance for helpfulness prediction?
- How much the review helpfulness of product reviews varies as a function of search and experience product type?
- Which type of indicators (review, reviewer and product) strongly drives the online review helpfulness (quality)?
- What are the set of influential features among review, reviewer and product type for helpfulness?

The objective of this article is to explore the influential indicators of review content, reviewer and product type that strongly contribute to helpfulness of product review. The impact of each type of indicators on perceived helpfulness is examined by utilizing two real-life Amazon product review datasets. Six popular ML algorithms are trained and tested using proposed features and evaluated by three evaluation metrics. The effect of product type is also examined for review helpfulness. An effective prediction model is constructed for online review helpfulness using deep neural network method. The findings of the current study have added contributions to the prior studies by looking further into type of review content and reviewer indicators and their contribution toward helpfulness. These findings also extends the work done in prior studies (Mudambi and Schuff 2010; Lee and Choeh 2014) by intensively investigating textual and reviewer characteristics. Major contributions are:

1. This research proposes three types of features, namely review content, reviewer and product for helpfulness prediction of online reviews using two Amazon datasets. In addition, the effect of product type is also investigated for review helpfulness.
2. We are the first ones that utilize deep neural network for building a review helpfulness prediction model.
3. It has been observed from the experiments that polarity of review title, sentiment and polarity of review text, cosine similarity between review text and product title; reviewer production time and active since; number of words in product title and number of questions answered are the most influential indicators.
4. This research enables online managers, merchants and retailers to better enlist product reviews for online customers by minimizing the processing costs.

The research article is organized as follows. Section 3 presents the related work followed by the Sect. 4, which presents the proposed methodology of this article. Section 5 conducts analysis of various experiments and results, and finally discussions. Subsequently, Sect. 6 presents the implications and Sect. 7 presents concluding remarks and outline directions for future work.

3 Literature review

3.1 Qualitative factors and review helpfulness

Liu and Cao (2007) used readability, subjectivity and informativeness characteristics to identify the low- and high-quality reviews from large volume of reviews. Then, Zhang and Varadarajan (2006) developed a regression model by using part of speech terms along with subjectivity, lexical and similarity as determinants. The proposed model predicts the utility of customer reviews. Later, another helpfulness predictive model is presented by Mudambi and Schuff (2010) and identifies the significant predictors for review helpfulness. Then, Huang and Yen (2013) replicated their research and attained only 15% explanatory power. Later, impact of reviewers, and qualitative and quantitative characteristics of reviews are explored by Huang et al. (2015) to examine their impact on helpfulness. Reviewer cumulative helpfulness, experience and impact are used as reviewer features. The experiments demonstrate that reviewer experience has a varying effect, and word count with a certain threshold is effective for the helpfulness of user reviews.

Recently, qualitative characteristics of review content are investigated by Agnihotri et al. (2016) to examine their influence on helpfulness prediction of online reviews. The results demonstrated that reviewer experience has a moderating role which affects the trust of online customers. Later, a conceptual helpfulness predictive model is built by Qazi et al. (2016). The qualitative and quantitative characteristics of reviewer are utilized as features. The findings reveal that review types and average number of concepts per sentence have varying effects on helpfulness of users' reviews. More recently, an order effect of reviews on helpfulness is investigated in the prospective of social influence (Zhou and Guo 2017). It is the first attempt, and results demonstrate that review order is negatively related to review helpfulness and that negative influence is inversely proportional to reviewer expertise.

3.2 Linguistics and sentiment factors for review helpfulness

Lexical, structural and meta-data indicators of review text are utilized by Kim and Pantel (2006) to predict review helpfulness. The results reveal that review length, sentiment and its unigrams are the most effective indicators. Then, a model is designed by Hong and Lu (2012), to classify the product reviews. Need fulfillment, text reliability and review sentiment are utilized as features. The classification is performed using SVM method, and it outperformed the work of previous studies (Kim and Pantel

2006; Liu and Cao 2007). Later, a binary helpfulness prediction model is developed by (Krishnamoorthy 2015). Author derived an algorithm to extract linguistic characteristics, and results demonstrated that the proposed features deliver the best performance. Ullah et al. (2015) presented an approach that examined the influence of emotion characteristics on review helpfulness. The findings reveal that contents with negative emotions have no effect and contents with positive emotions have positive influence on review helpfulness. Later, the influence of product type and review sentiment on review helpfulness is explored by Chua and Banerjee (2016). The results demonstrated that review helpfulness varies across information quality as a function of review sentiment and product type. Then, Ullah et al. (2016) investigated the impact of emotions of review text on review helpfulness. The findings reveal that there are different influences of emotions on helpfulness across experience and search goods. Another helpfulness prediction model is built by Malik and Hussain (2017) and utilized four discrete positive and four negative emotions. Authors found that positive emotions have more influence on review helpfulness than negative emotions. Recently, impact of linguistic characteristics is investigated by Malik and Iqbal (2018) to determine their contributions for helpfulness prediction. They concluded that noun and proposition are most effective indicators for review helpfulness.

3.3 Quantitative factors for review helpfulness

An unsupervised method is proposed by Tsur and Rapoport (2009) to rank the book reviews on the basis of helpfulness score. Author designed a lexicon which contains virtual core reviews and dominated terms. The ranking of reviews is computed on the basis of similarity between review text and virtual core review. Later, Lee and Choeh (2014) applied multilayer perceptron neural network to build a review helpfulness predictive model. Textual, meta-data, reviewer and product characteristics are utilized, and MLP neural network outperformed the linear regression model. Recently, confidence interval and helpfulness distribution indicators are utilized by Zhang et al. (2014) to predict review helpfulness. The synthetic and real datasets are utilized in the experiments to investigate the influence of the proposed features. The reviewer recency, frequency and monetary value features are introduced by Ngo-Ye and Sinha (2014) to effectively predict the review helpfulness. The results reveal that hybrid feature combination delivers the best performance. But authors did not consider meta-data, subjectivity and readability which are experimentally verified by prior studies (Kim and Pantel 2006; Liu and Cao 2007; Ghose and Ipeirotis 2011).

Recently, reviewer and review text characteristics are combined to construct a regression model for review helpfulness prediction (Liu and Park 2015). The findings indicate that reviewer reputation, expertise and identity, valence and readability of review text are effective features. Then, (Chen et al. 2015) investigated the impact of existing votes, reviewer and review indicators on helpfulness. The experiments reveal that review valence, review votes and reviewer indicators are strong predictors of helpfulness of product reviews. Later, a multilevel regression algorithm is applied to explore the influence of valence consistency on helpfulness (Quaschnig et al. 2015). The impact of nearby reviews is investigated instead of individual reviews. The findings indicate that consistent reviews attract more helpful votes.

A conceptual model is built by Chen (2016), in which relationship between helpfulness of product reviews and review sidedness is examined. The experiments found that in the presence of expert writing reviewers for search goods, two-sided reviews attract more helpful votes than one-sided reviews. Later, impact of reviewer image, review valence, depth and equivocality for helpfulness prediction are investigated by Karimi and Wang (2016). The findings indicate that reviewer image is an important determinant. Yang and Shin (2016) conducted an exploratory study and examined the relationships between review helpfulness and review photograph, review rating, length, reviewer location and level and helpful votes. The findings indicate that reviewer helpful votes and review rating are strong determinants. However, the social effect of reviewers may increase the predictive accuracy of review helpfulness.

Recently, a MSP model tree is applied to predict the helpfulness of hotel reviews by Hu and Chen (2016). Interaction effect between review rating and hotel stars and visibility are utilized as features. The results reveal that visibility of review is the effective predictor, and a relationship exists between rating and hotel stars. Then, Xiang and Sun (2016) investigated the influence of multi-typed indicators on helpfulness and built a Berlo index communication system. Later, Gao et al. (2017) utilized consistency of reviewer's pattern of rating over time and predictability as determinants to examine their impact on review helpfulness. The findings indicate that reviews will attract more helpful votes in future if they hold higher absolute bias in rating. In addition, rating behavior of reviewer is consistent along products and time. Prior studies used textual and non-textual indicators and regression methods (Cao et al. 2011; Ghose and Ipeirotis 2011; Pan and Zhang 2011; Korfiatis et al. 2012; Chua and Banerjee 2015).

The influence of reviewer identity, reputation, review depth and moderating effect of product type for helpfulness prediction is investigated by Lee and Choeh (2016). The

findings reveal that review depth and extremity, and reviewer reputation are strong predictors for review helpfulness across search goods. Several popular ML algorithms are applied to construct an effective helpfulness prediction model by Singh et al. (2017). Review subjectivity, polarity, entropy and reading ease are utilized as features. Recently, Ngo-Ye et al. (2017) utilized regression method to evaluate the influence of script analysis for review helpfulness. The results indicate that proposed helpfulness model presents better performance with reduced training time. More recently, Malik and Hussain (2018) investigated the influence of reviewer and review-type variables on helpfulness of reviews. The results indicate that reviews are more effective than reviewer indicators. We believe that these results are relevant in the context of the current research. The use of influential reviewer, review and product indicators offers better predictive results.

4 Research methodology

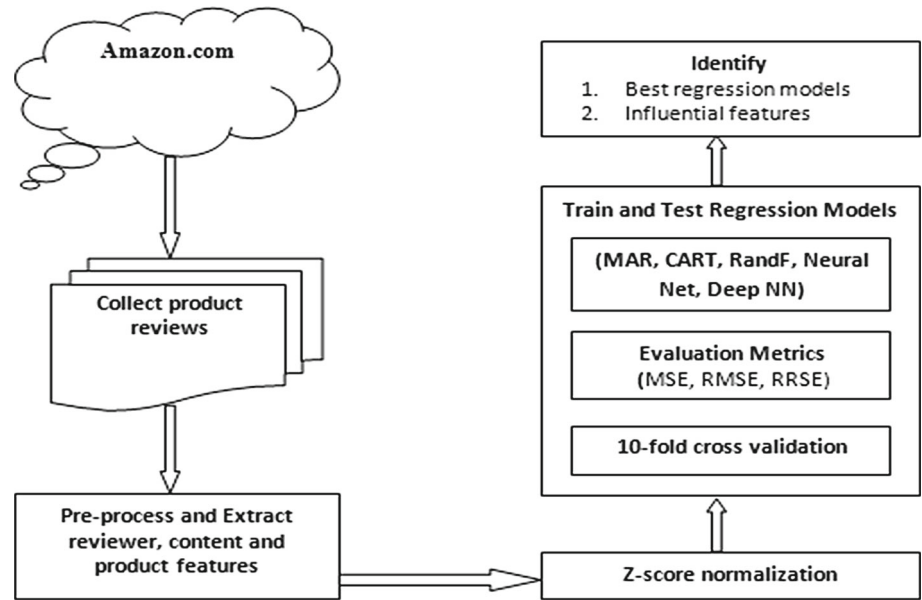
The methodology adapted in this research is presented in Fig. 1. First, reviews of thousands of products are collected from Amazon.com. Second, these product reviews are further preprocessed, and reviewer, content and product-type features are computed. Third, z-score normalization is performed for all computed features to scale the features values. Fourth, five regression models including multivariate adaptive regression (MAR), classification and regression tree (CART), random forest (RandF), neural network (Neural Net) and deep neural network (Deep NN) are chosen to construct the predictive models for review helpfulness. These regression models are evaluated using mean square error (MSE), root mean square error (RMSE) and root relative square error (RRSE)-based evaluation metrics. Tenfold cross-validation method is utilized in all sorts of experiments. Fifth, outperformed learned model and set of influential features are identified in order to build an effective review helpfulness prediction model.

In this study, the examined case is the regression problem in which the dependent variable is review helpfulness. It is defined as the ratio of helpful votes to the total votes received by a review. The independent variables or estimators may be classified into three categories: reviewer, review content and product characteristics. The descriptions of these estimators are discussed in the remaining part of this section.

4.1 Estimators

This section is intending to introduce proposed review content, reviewer and product-type features in addition to

Fig. 1 Research method



the state-of-the-art baseline features for helpfulness prediction of reviews. The features are: (1) review content features, (2) reviewer features, (3) product features, (4) linguistic features, (5) readability features and (6) visibility features. The following subsections will comprehensively present the descriptions of each type of features.

4.1.1 Review content features

In the current research, we introduced novel indicators of review content in order to predict helpfulness of reviews effectively. Prior studies have proved the importance of review textual/content characteristics for helpfulness analysis (Lee and Choeh 2014; Chua and Banerjee 2016). The proposed features are: (1) NumComments (number of comments posted on a review), (2) CosineSimi (cosine similarity of review text and product title), (3) SentiText (sentiment of review text), (4) PolText (polarity of review text), (5) NumWord (number of words in review title), (6) SentiTitle (sentiment of review title) and (7) PolTitle (polarity of review title).

CosineSimi: Cosine similarity measures the angle-based similarity of product title and review text. The objective of this indicator is to find in what percentage product title and review text are similar. In particular, higher CosineSimi means both (review text and product title) texts are more similar. This shows that review text contains more words which are similar to the words used in product title. In other words, product title will be the representative of review text (for higher CosineSimi) and that review definitely attracts more helpful votes. Mathematically, it is computed as:

$$\text{Cos}(x, y) = \frac{x \cdot y}{xy} = \frac{\sum_{i=0}^{n-1} x_i y_i}{\sqrt{\sum_{i=0}^{n-1} x_i^2} \sqrt{\sum_{i=0}^{n-1} y_i^2}} \quad (1)$$

where x and y are two vectors, and number of comments posted on a review text can be obtained from Amazon.com.

SentiText, PolText, SentiTitle, PolTitle: The sentiment and polarity of review title and review text are computed by using SentiStrength software (Thelwall et al. 2010). This app is able to analyze positive and negative emotions and booster words in the text. The objective of sentiment analysis in this research is to determine overall contextual polarity or emotional reaction in the review text and title. The objective of polarity computation of review text/title is to classify the expressed opinion into positive, negative or neutral. The positive sentiment ranges from 1 to 5, and negative sentiment ranges from -1 to -5, where 1 represents least positive and 5 represents extremely positive. Similarly, -1 indicates least negative and -5 indicates extremely negative. The sentiment and polarity values of review title and review text are computed by utilizing the method (Stieglitz and Dang-Xuan 2013). Mathematically,

$$\text{PolTitle} = \text{PolText} = \text{Positive Sentiment} + \text{Negative Sentiment} \quad (2)$$

$$\text{SentiTitle} = \text{SentiText} = \frac{(\text{Positive Sentiment} - \text{Negative Sentiment})}{-2} \quad (3)$$

4.1.2 Reviewer features

Reviewer features are also influential features for review helpfulness prediction (Lee and Choeh 2014). This research also proposed two indicators of reviewer for helpfulness prediction analysis. The features are: (1) *AUSince* (Length of a time since the first review posted by the reviewer) and (2) *ProdTime* (Length of a time between the first review and the recent review posted by the reviewer).

AUSince Prior literature reveals that with the passage of time, a number of helpful reviews get accumulated. Similarly, the number of days since the first review posted by the reviewer is an effective determinant for review helpfulness prediction. We expect positive correlation on temporal recency; a reviewer with longer active time may write more helpful reviews for products.

ProdTime Prior research demonstrated that temporal dimension has a significant role for Author citation prediction task (Daud et al. 2015). Production time (*ProdTime*) of reviewer is computed as the number of days between the first review and the recent review posted by a particular reviewer. This indicator shows that how much long a reviewer is actively participating in writing reviews for products. In the same way, longer production time of a reviewer represents that there is large collection of potential reviews written by the reviewer.

4.1.3 Product features

This study proposed five product-related indicators to investigate the impact of product determinants on review helpfulness. The type of product determinants is also utilized by prior studies (Ghose and Ipeirotis 2011; Lee and Choeh 2014). The proposed features are: (1) *PerPos* (percentage of positive reviews), (2) *PerCrit* (percentage of critical reviews), (3) *NumWordsP* (number of words in product title), (4) *NumQ* (number of questions answered) and (5) *PScore* (potential score of a product). The total number of questions answered for each product is obtained from Amazon.com. The percentages of positive and critical reviews are computed as follows:

$$\text{PerPos} = (\text{Number of positive reviews} * 100) / \text{Total reviews} \quad (4)$$

$$\text{PerCrit} = (\text{Number of critical reviews} * 100) / \text{Total reviews} \quad (5)$$

where total number of positive and critical reviews of the product is obtained from Amazon.com. Amazon.com provides the count of total positive and total critical reviews for each product.

PerPos*, *PerCrit Percentage of positive reviews and critical reviews is important indicators for review helpfulness. A product with large count of *PerPos* indicates the evidence of users' satisfaction.

PScore It is the potential score of a product. The large score of a product indicates that although there is less volume of reviews posted for the product, all reviews attracted high rating. Alternatively, the product which receives large review rating and less number of reviews will get large potential score because small value of denominator (number of reviews) will generate higher potential score for the product.

4.1.4 Linguistic features

Prior studies proved that linguistic indicators of review content are also influential determinants that aid in identification of helpful reviews. This study utilizes state-of-the-art linguistic features (Singh et al. 2017), such as '**Nouns**,' '**Verbs**,' '**Adverbs**' and '**Adjectives**,' to compare the performance of the proposed features with state-of-the-art baseline features for review helpfulness prediction. The values of these features are computed by counting their respective percentages in the review text.

4.1.5 Readability features

Readability of review text is established to be an effective indicator for online review helpfulness prediction (Hu and Chen 2016). The set of reviews with high readability is definitely to be read by users and attracted more helpful votes. This study used the following grade-level readability indexes for comparisons:

1. SMOG (Simple Measure of Gobbledygook)
2. ARI (Automated Readability Index),
3. GFI (Gunning Fog Index),
4. FKGL (Flesch–Kincaid Grade Level)
5. CLI (Coleman–Liau Index)
6. FKRE (Flesch–Kincaid Reading Ease).

4.1.6 Visibility features

Prior studies demonstrated that review visibility determinants are important features to predict review helpfulness. This article also utilized six visibility features for performance comparisons (Lee and Choeh 2014). The features are:

1. Rating (Rating of the review)
2. EDays (Elapsed days since the posting date)
3. SentiRating (Sentiment of review in terms of rating)
4. RExtrem (Review Extremity).

5. Len_words (Length of a review in words)
6. Len_Sentences (Length of a review in sentences)

4.2 Data collection and pre-processing

This research utilized two real-life Amazon product review datasets for the experimental setup. The first dataset (DS1) contains reviews from 21 different product categories, as listed in Table 1. This dataset is publicly available, named as multi-domain sentiment analysis dataset (Blitzer and Dredze 2007). It initially contains 150,000 product reviews. The second dataset (DS2) is obtained by crawling reviews from 34 different product categories of Amazon.com. It contains more recent product reviews from novel 34 different product categories. Initially, the second dataset contains 40,800 reviews. There are 40 different product categories available on Amazon.com. The second dataset is novel because it contains reviews from a majority of product categories of Amazon.com. The number of different products in the second dataset is 3360, and only those products are considered which comprise the top-10 best sellers from each category. The list of product categories for both datasets is presented in Table 1.

Then, data cleaning process (Liu and Huang 2008) is applied to further pre-process both datasets and finally z-score normalization is applied to normalize the data. The steps of data cleaning are: (1) Duplicate reviews are identified and removed, (2) reviews with blank text are also removed from both datasets, and (3) reviews which have large helpful votes but very low total votes are less useful. Therefore, only those reviews are selected which have at least 10 total votes. Finally, we have 109,357 reviews and 32,434 reviews in the first and second datasets, respectively (Table 1). The complete steps for data pre-processing are demonstrated in Fig. 2. Then, the frequencies of various helpfulness scores for both datasets are drawn in Fig. 3. For this problem, helpfulness is the target value and it varies from 0 to 1. Helpfulness is computed as the ratio of helpful votes to the total votes attracted by a review, and that is

why the lower bound and upper bound for helpfulness is 0 and 1. All the prior researches used this formula for helpfulness (ranges from 0 to 1). Frequencies are plotted along the y-axis, and range of helpfulness is plotted along the x-axis and it shows the frequencies of reviews at various helpfulness scores. Both datasets hold large volume of reviews at helpfulness range from 0.9 to 1. This indicates that density of helpfulness for both datasets is skewed toward the right.

4.3 ML models and evaluation metrics

Popular machine learning models are selected in this research to build a helpfulness prediction model using reviewer, review content and product-type features. The models are (1) MAR, (2) CART, (3) RandF, (4) Neural Net and (5) Deep NN. We are the first ones that are using Deep NN method to build an effective helpfulness predictive model. The maximum size of feature matrix is $N * 28$, where 28 means maximum number of features and N is the number of instances. Standard tenfold cross-validation approach is used in all sorts of experiments, and MSE-, RMSE- and RRSE-based evaluation metrics are used to evaluate the performance of various regression models. R language is utilized for models' training and testing.

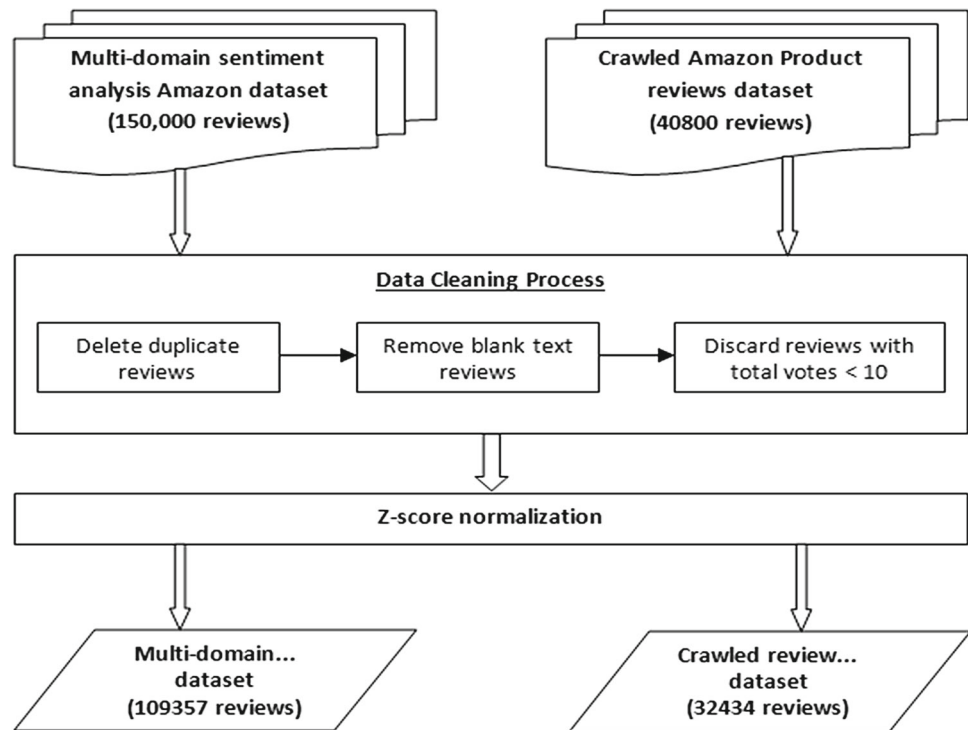
4.3.1 Deep neural network

This research utilized five-layer feedforward artificial neural network with stochastic gradient descent and back-propagation technique as a training function for review helpfulness prediction task since deep neural networks are able to learn complex nonlinear relationships between input and output variables using past samples. Previous studies showed that they have proved their popularity due to their effective performance in diverse fields. This research use feedforward Deep NN to perform the prediction task of review helpfulness, and five layers are selected in the internal architecture. The data instances are

Table 1 Dataset

Dataset	No. of reviews	Product categories
DS1	109,357	Apparel, automotive, baby, beauty, camera and photograph, cell phones and service, computer and video games, electronics, gourmet food, grocery, health and personal care, jewelry and watches, kitchen and house wares, magazines, musical instruments, office products, outdoor living, software, tools and hardware, toys and games, video
DS2	32,434	Camera and photograph, cell phones and accessories, clothing, appliances, books, App Store for android, arts, crafts and sewing, computers and accessories, electronics, gift cards, beauty, industrial and scientific, automotive, baby, grocery and gourmet food, health and personal care, home and kitchen, home improvement, jewelry, kindle store, kitchen and dining, MP3 downloads, magazines, movies and TV, music, musical instruments, office products, patio, lawn and garden, pet supplies, shoes, software, sports and outdoors, toys and games, video games, watches

Fig. 2 Data pre-processing flowchart



applied in the input layer. Then, the output of input layer is applied into three hidden layers (hidden1, hidden2, hidden3) step by step. Then, outcome of hidden3 layer is applied to the output layer in which output of the Deep NN is computed. The graphical representation of architecture of Deep NN is presented in Fig. 4.

There is one neuron in the output layer, and tanh is utilized as an activation function that is bounded by (0, 1), where Net is the sum of weighted input. There are 28 neurons configured in the input layer, and same numbers of neurons are configured at hidden1, hidden2 and hidden3 layers to moderately perform the tasks of both training and testing. Each neuron in a layer is connected in the forward direction to each neuron in the next layer. All initial weights are configured at 0.3. Table 2 presents the necessary parameters of Deep NN.

5 Experimental results

In this study, we conducted various experiments to examine the significance of proposed features on customer review helpfulness and to compare the performance of the proposed features with state-of-the-art baseline features using five machine learning regression models. The performances of the helpfulness predictive models are compared using popular MSE, RMSE and RRSE evaluation metrics. The experiments are helpfulness prediction analysis, feature analysis, feature-importance analysis and

analysis of search and experience product types on review helpfulness.

5.1 Prediction analysis

This set of experiments built prediction models for review helpfulness by using five regression techniques and standard tenfold cross-validation method. Five regression methods are trained and tested on two datasets using proposed hybrid set of features (combination of review, reviewer and product features). The performance of applied regression models is compared using MSE, RMSE and RRSE evaluation metrics. Figure 5 demonstrates the predictive performance of five regression models using MSE-based metric. Due to powerful predictive capability, Deep NN outperforms and presents the lowest mean square error value as compared to other four regression models for two datasets. However, DS2 dataset presents better results as compared to DS1 dataset. Similarly, RMSE- and RRSE-based predictive performance for review helpfulness is presented in Fig. 6. Deep NN presents least errors as compared to other state-of-the-art regression models using hybrid combination of features on two datasets.

The experimental results reveal the outperformance of Deep NN for both datasets using proposed hybrid combination of features. DS1 dataset presents 0.07 MSE, 0.261 RMSE and 0.566 RRSE. However, DS2 dataset exhibits superior performance by delivering 0.06 MSE, 0.237 RMSE and 0.478 RRSE. In addition, the least values of

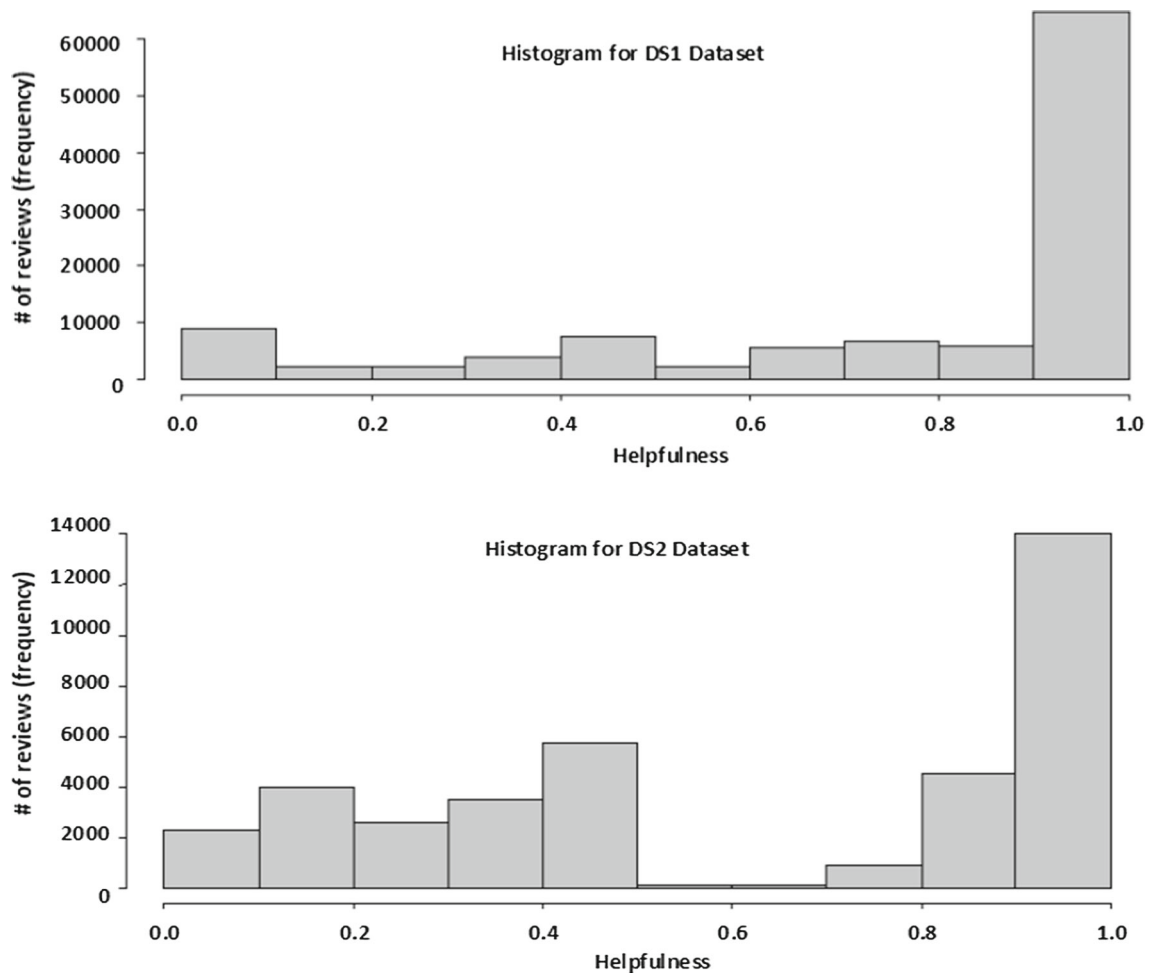


Fig. 3 Histogram of helpfulness for both datasets

Fig. 4 Deep neural network architecture diagram (Lozano-Diez et al. 2017)

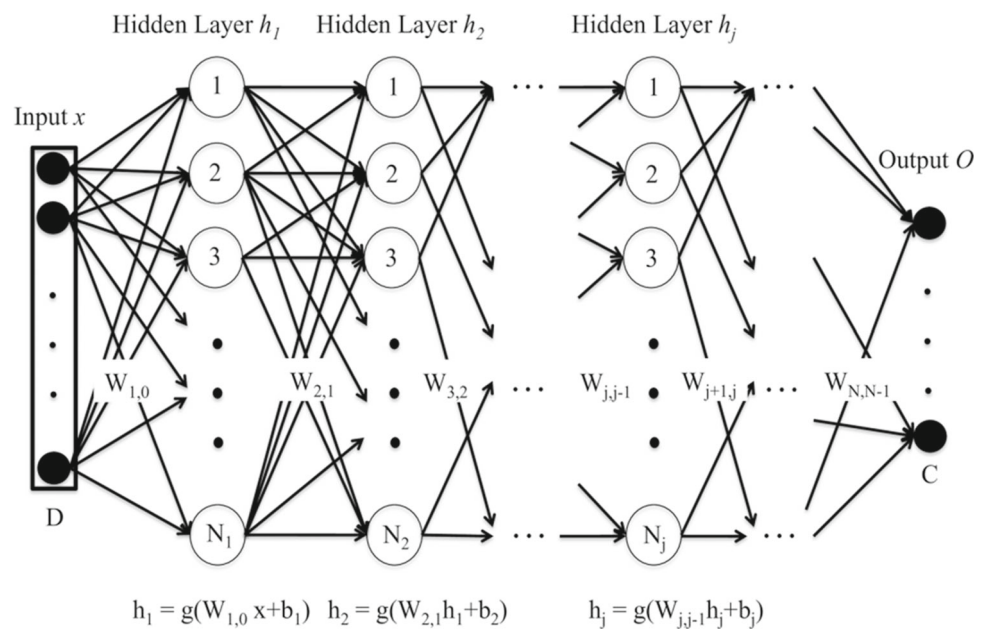


Table 2 Deep NN parameters

Parameters	Value
Distribution function	Gaussian
Initial weights	0.3
Loss function	Cross-entropy
Momentum start	0.0005
Momentum rate annealing	1.0E−6
Learning rate	0.05
Learning rate decay	0.5
Activation function	$\text{Tanh} = f(\text{NET}) = \frac{e^{\text{NET}} - e^{-\text{NET}}}{e^{\text{NET}} + e^{-\text{NET}}}$
Epochs	100

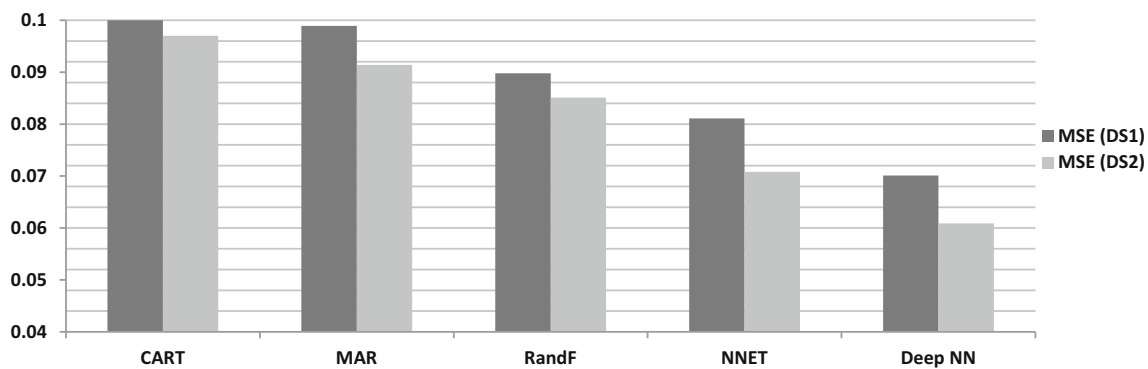
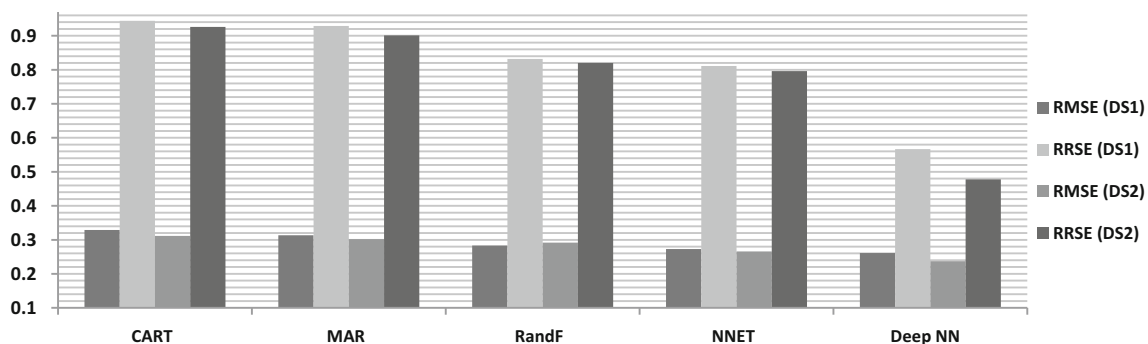
MSE, RMSE and RRSE for both datasets produced by Deep NN indicate that it delivers promising performance and results prove the utility of the proposed features (review type, reviewer and product) for review helpfulness prediction. The Deep NN is utilized for the remaining set of experiments because it delivered the best results.

5.2 Feature analysis

This section conducted the experiments to examine the significance of each type of characteristics to predict the

product review helpfulness using two datasets. Deep NN is utilized as a regression model to conduct these experiments. There are six types of features utilized in this research for prediction analysis and comparisons. The proposed features are product, reviewer and review types, whereas type of linguistic, readability and visibility features are state-of-the-art baseline features. The three types of proposed features (product, reviewer and review) are combined as a hybrid proposed set, and the remaining three (linguistic, readability and review visibility) are combined as a baseline set.

The experimental results (Fig. 7) show that by considering individual type of features (standalone model) for helpfulness analysis using two datasets, the best predictive performance is obtained by review features as compared to other type of features. This exhibits that review content is the most influential indicator among the six types of features for both datasets. The reviewer-type and review visibility-type features deliver comparable performance, but better than product, readability and linguistics-type features using both datasets (Fig. 7). This shows that reviewer features are less significant than review content, but better than other features to predict the helpfulness of product reviews. In addition, type of product features presents the lower predictive performance.

**Fig. 5** MSE-based predictive performance using hybrid set of features**Fig. 6** RMSE- and RRSE-based predictive performance using hybrid set of features

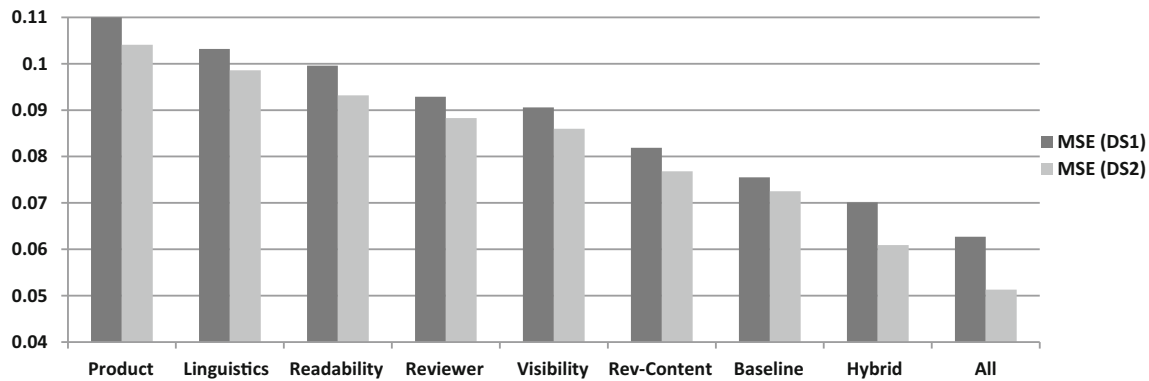


Fig. 7 Feature-wise performance analysis

It is evident from the results that hybrid set of proposed features presents the least MSE values as compared to baseline set of features for online helpfulness prediction using two datasets. However, DS2 delivers better performance as compared to DS1 for all set of experiments. The results demonstrate that the proposed features are more influential than baseline features for review helpfulness prediction when proposed and baseline features are combined in their respective sets. In the same way, we get the best performance by combining all features, as shown in Fig. 7. Previous studies revealed that the review and reviewer characteristics are the most important determinants for helpfulness prediction (Korfiatis et al. 2012; Lee and Choeh 2014). Our research supports the findings of prior studies and demonstrates the outperformance of the proposed review and reviewer features with promising results.

5.3 Feature importance

In this section, a number of experiments are simulated to examine the importance of each proposed feature related to product, reviewer and review types to predict helpfulness of product reviews using both datasets (DS1 and DS2). The objectives of these experiments are to investigate how much each indicator contributes to review helpfulness and which indicator contributes the most.

The importance of the proposed product, reviewer and review features using DS1 dataset is presented in Fig. 8. The significance is evaluated on the basis of mean square error. The experiment demonstrates that review features are the most important indicators. Polarity of title is the most influential feature of review type which indicates that more positive or negative titles of reviews attract more readerships, and therefore receive more helpful votes. Similarly, sentiment and polarity of review text are the next influential predictors for review helpfulness. This shows that reviews which have high positive or negative sentiment

scores attract more helpful votes. Cosine similarity is the next influential feature of review type, as shown in Fig. 8. If the value of cosine similarity is higher, then more helpful votes are attracted by that particular review. Number of words in title and sentiment of title have similar importance for review helpfulness prediction. However, number of comments is the least important feature among review type for review helpfulness prediction.

Reviewer production time is the most significant reviewer-type predictor for helpfulness prediction of product reviews, as demonstrated in Fig. 8. However, it is less effective than polarity of title and sentiment of review features. The significance of production time indicates that reviews posted by the potential reviewers who have large production time attract more helpful votes as large production time indicates that reviewer has a lot of experience in writing influential reviews. Reviewer active since is the next effective reviewer-type indicator for review helpfulness prediction. It is evident from Fig. 8 that reviewer active since and cosine similarity have almost similar importance. However, the experiments reveal that product-type features are least effective than reviewer and review-type features (Fig. 8). Number of words in product title is the most effective, and percentage of critical reviews is the least effective indicators of product type.

We obtained better results from experiments conducted to compute the importance of proposed features using DS2 dataset. Mean square error-based metric is utilized to analyze the performance of each indicator, as shown in Fig. 9. The type of review features is again the most influential than reviewer and product-type features. Polarity of title outperforms and obtains the lowest MSE value as compared to MSE obtained by other indicators using DS2 dataset. All proposed features achieved low MSE values using DS2 dataset as compared to DS1 dataset. The next influential feature is polarity of review text in contrast to sentiment of text which is identified as second influential by using DS1 dataset. All other features of review type

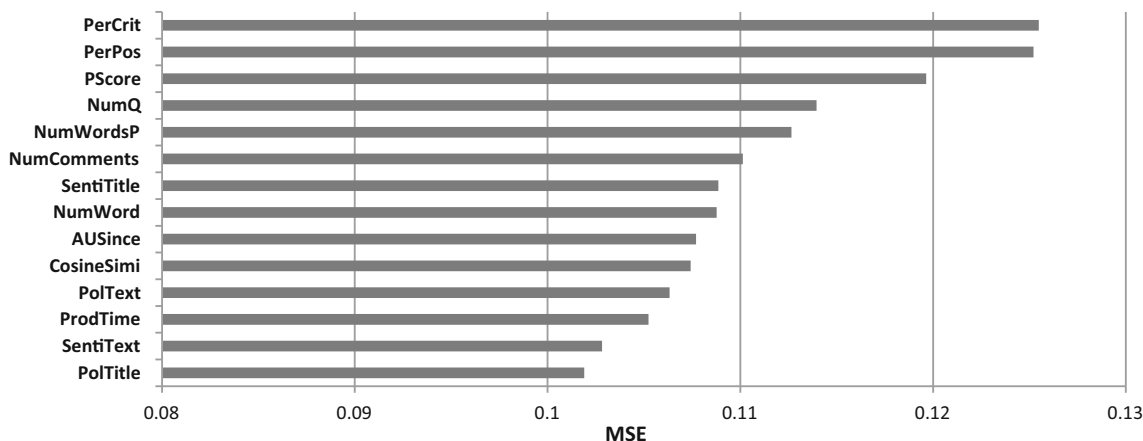


Fig. 8 Feature importance using DS1 dataset

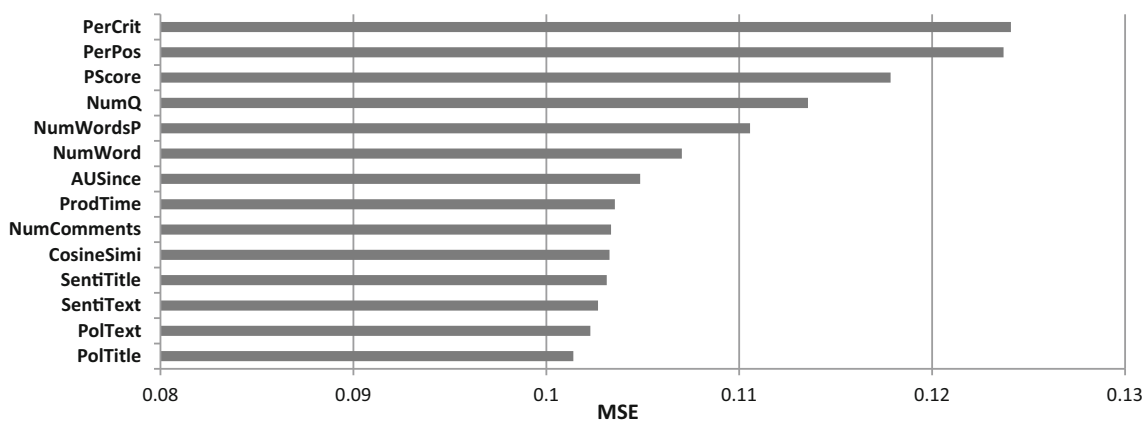


Fig. 9 Feature importance using DS2 dataset

obtain the same rankings by using DS1 dataset except number of words in title, which is identified as least important by using DS2 dataset (Fig. 9). Reviewer production time and number of words in product title are the most important among reviewer and product-type features.

5.4 Influence of product type on helpfulness

To investigate the impact of product type on helpfulness of product reviews using proposed features, an experiment is conducted using DS1 and DS2 datasets. It is evident from prior studies (Nelson 1970; Huang and Yen 2013) that the products could be classified into two types, i.e., search and experience goods, and helpfulness of online reviews varies along both product types. From both datasets, search and experience products’ reviews are separated into two sets and Deep NN is used as the regression model. The aim of the experiment is to examine the influence of the proposed features of users’ reviews that belong to search and experience goods on helpfulness prediction. The findings indicate that the proposed features demonstrate strong

Table 3 Impact of product type on helpfulness using DS1 and DS2

Features	Search		Experience	
	MSE (DS1)	MSE (DS2)	MSE (DS1)	MSE (DS2)
Product	0.1025	0.0993	0.1172	0.1127
Linguistics	0.092	0.0887	0.1091	0.1049
Readability	0.0893	0.0862	0.1047	0.1008
Reviewer	0.0856	0.0717	0.0999	0.0938
Visibility	0.0837	0.0723	0.0993	0.0931
Rev-Content	0.0723	0.0604	0.0906	0.0824
Baseline	0.0711	0.0596	0.0808	0.0789
Hybrid	0.0601	0.0531	0.0737	0.071
All	0.0502	0.0482	0.0702	0.0642

relationship to helpfulness using search goods in contrast to experience goods for both datasets (Table 3). Reviews of search goods definitely contain useful information about significant product attributes and other aspects, but reviews

of experience goods are comprised of various user experiences about products. Similarly, hybrid set of proposed features using search goods presents better predictive results as compared to experience goods for both datasets. In addition, performance of various features types and their combinations in case of search goods is better than results shown in Fig. 7. We obtained the best predictive performance in case of search goods when all features are combined.

5.5 Discussion

The findings of current research demonstrate strong empirical support to our research contributions. Identification of influential indicators of helpfulness of product reviews has attracted much interest in the prior studies. However, key insights may become indistinct without consideration of strong features. The contributions of this study are robust and consistent across two real-life Amazon datasets, review, reviewer and product-type indicators and three evaluation metrics. This research proposed a review helpfulness prediction model that utilizes review, reviewer and product-type indicators using Deep NN regression model. The types of visibility, readability, linguistic features are utilized as baseline for comparisons. The proposed helpfulness predictive model is very significant to identify the list of helpful reviews that will help online users in making purchasing decisions. Deep NN model is established to be more effective than other four regression models (Figs. 5, 6). The results demonstrate that review characteristics are the most influential, and the best performance (MSE of 0.07 and 0.06 using DS1 and DS2 datasets) is obtained using hybrid set of proposed features (Fig. 7).

We found Deep NN method as the best ML model in the experiments (Figs. 5, 6). It delivers minimum MSE as compared to MAR, CART, RandF and Neural Net ML methods. It is the first study that utilizes Deep NN to construct a predictive model for review helpfulness. The reviewer and review-type indicators have proved their significance as the strong predictors to improve helpfulness predictive accuracy for both datasets. As a standalone model, review indicator shows better performance as compared to other types of indicators (Fig. 7). In addition, hybrid combination of the proposed features delivers better performance as compared to hybrid combination of baseline features. The combination of all features demonstrates the best performance. The experiments (Figs. 8, 9) conducted to investigate the importance of proposed features reveal that polarity of review title is the most important indicator of review type, whereas sentiment and polarity of review text and cosine similarity also show strong relationships to helpfulness of product reviews.

Similarly, reviewer production time demonstrates strong relationship to review helpfulness among reviewer-type features. Although product indicators are less effective, however, a number of words in product title and number of question answered are the strong product-type indicators. The influence of proposed features on helpfulness of user reviews in the presence of product types (search and experience) is also investigated (Table 3), and conclusions are drawn. It has been evaluated that proposed review, reviewer and product-type indicators are more influential using search product type as compared to experience product type. In addition, review indicator presents the best predictive performance for search goods as a standalone model. The set of experiments and their findings provide the evidence of utility of the proposed characteristics.

This research also has several limitations. First, both datasets are designed by using Amazon product reviews from popular online e-commerce retailers (Amazon.com) that enable online users to leave their opinions and suggestions about popular products. Similarly, Rakuten.com and Newegg.com e-commerce merchants also supported these kinds of accessories. Moreover, physical stores such as Wal-Mart and best-buy enable users to leave their opinions and feedbacks. However, as Amazon is one of the popular retailers, utilizing Amazon dataset to generalize to overall market could be biased. That is why online users should not generalize the results beyond the intended context. However, the volumes of trades being processed and size of Amazon organization have made it one of the largest online retailers. Data sample taken from Amazon.com still represents a large proportion of the true online retailing population. Future studies can include more products and different brands in order to further enhance findings presented in this article.

Second, as users' reviews have also been shown to influence sales, different market participants have already begun to make public very positive online reviews to boost the sales of their offered products or extremely negative reviews to reduce the turnover of their competitors. As follows, an analysis of reviews posted on the Internet is accompanied by the risk that such fake reviews are included in the dataset, which might bias the results. Nevertheless, we analyze different products, so a manipulation of a single product or service would have only a minor impact on the results. Additionally, because the different products are best sellers and are thus discussed within a large number of reviews, a potential manipulator would need to publish a large number of fake reviews, which makes manipulation time-consuming and, consequently, less probable.

Finally, another limitation is the self-selection bias in our dataset, because we only included review posters. The opinions from the non-posting consumers who brought a

particular product from current E-commerce sites were not analyzed in our research. This is the common drawback for studies on related topic that collected secondary data from online product review websites. Therefore, further empirical researches are encouraged to explore the issue of review helpfulness using other complementary research methods, such as survey or eye-tracking experiments.

6 Implications

The findings of the current study have several implications in this domain. Prior developments for online review helpfulness prediction have adopted distinctive approaches and produced worthy, but diverse findings, and sometimes drew inconsistent conclusions. The objective of current research is to build an effective predictive model for online review helpfulness. We are the first ones that built a helpfulness predictive model based on Deep NN regression method. Deep NN models are capable of estimating the complex relationships between features because they can capture nonlinear relationships in the data. These models do not depend on assumptions of distribution of error terms and are usually better than statistical approaches. The proposed methodology also utilizes the novel combination of the proposed features, i.e., review, reviewer and product types. The experimental evaluations verified the significance of proposed methodology.

Generally, this research presents that helpfulness of product reviews is a complex concept and it is a continuing effort demonstrated by the prior studies. Polarity of title, sentiment and polarity of review text and cosine similarity have strong influence on helpfulness rating. Reviewers who wish to be helpful must remember that there should be consistency in review quality. The results also exhibit that influential reviewers posted large volume of reviews than average reviewers. Alternatively, large volume of reviews was more likely to be written by the influential reviewers. This increases the chances of more reviews that attracted helpful votes by the customers. The current research extended the work of prior studies (Mudambi and Schuff 2010; Lee and Choeh 2014) by applying Deep NN with the combination of review, reviewer and product-type features. These developments could be effectively applied in multiple domains such as review summarization, recommendation, opinion mining and sentiment analysis.

Internet sellers and merchants provide the platforms for online customers to leave their opinions and feedbacks about products. The platforms also support online users to give votes for useful reviews. These opinions and feedbacks provided by the users about product evaluations (product usage experiences, etc.) increase the user trust and aid in the purchase decisions. Usually, customers encounter

the problem of review overload that requires high processing costs in decision-making process. The findings of current research enable the product retailers to embed necessary functionalities and to minimize the processing costs of user reviews for effective ranking.

The current research also draws multiple practical developments. The outcome of this research can be utilized to develop a smart review recommendation system for e-commerce products. Specifically, for a particular e-commerce website, whenever a user encounters a large volume of reviews for a specific product, the recommendation system can automatically identify helpful users' reviews on the basis of polarity of title, sentiment of text and number of words in product title indicators. It is a highly desirable requirement that online retailer should be capable of implementing a strong adaptive filtering because product users usually have short time to read all reviews of a particular product. Therefore, suggested recommendation system can facilitate users to easily gather important information of desired products by making better organization of online reviews.

7 Conclusions

This research investigates the influence of proposed review content, reviewer and product-type features on review helpfulness prediction. Five ML methods are utilized to build the helpfulness predictive models. We are the first ones that used Deep NN for helpfulness prediction problem. Seven review contents, two reviewers and five product features are proposed in this research. In addition, the state-of-the-art visibility, readability, linguistics features are also used for performance comparison and helpfulness prediction analysis. The experiments reveal that review content features deliver the best performance as a standalone model. However, reviewer and visibility features have comparable performance. In addition, hybrid proposed features (product, review and reviewer) deliver better performance as compared to hybrid baseline features. The variable importance computation demonstrates that polarity of review title, sentiment of review text, polarity of review text and cosine similarity between review title and text; production time and active since; number of words in product title and number of question answered are the most influential variables related to review content, reviewer and product-type features.

Several significant extensions can be made in future studies. Hybrid evolutionary algorithms can be applied to improve the prediction accuracy of helpfulness prediction model. One of the future extensions is to explore the use of influential features such as semantic and sentimental characteristics, reviewer identity and social features and

examine their impacts on review helpfulness. In addition, novel review and reviewer features can be applied in other domains such as to effectively rank the reviewers on the basis of influence score.

Funding This research did not receive any grant from funding agencies in the public, commercial or not-for-profit sectors.

Compliance with ethical standards

Conflict of interest The author declares that he/she has no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

- Agnihotri R, Dingus R et al (2016) Social media: influencing customer satisfaction in B2B sales. *Ind Mark Manag* 53:172–180
- Anderson M, Magruder J (2012) Learning from the crowd: regression discontinuity estimates of the effects of an online review database. *Econ J* 122(563):957–989
- Bertola F, Patti V (2016) Ontology-based affective models to organize artworks in the social semantic web. *Inf Process Manag* 52(1):139–162
- Blitzer J, Dredze M et al (2007) Biographies, bollywood, boom-boxes and blenders: domain adaptation for sentiment classification. In: *ACL*
- Cao Q, Duan W et al (2011) Exploring determinants of voting for the “helpfulness” of online user reviews: a text mining approach. *Decis Support Syst* 50(2):511–521
- Chen M-Y (2016) Can two-sided messages increase the helpfulness of online reviews? *Online Inf Rev* 40(3):316–332
- Chen Y, Xie J (2008) Online consumer review: word-of-mouth as a new element of marketing communication mix. *Manag Sci* 54(3):477–491
- Chen Y, Chai Y et al (2015) Analysis of review helpfulness based on consumer perspective. *Tsinghua Sci Technol* 20(3):293–305
- Chua AY, Banerjee S (2015) Understanding review helpfulness as a function of reviewer reputation, review rating, and review depth. *J Assoc Inf Sci Technol* 66(2):354–362
- Chua AY, Banerjee S (2016) Helpfulness of user-generated reviews as a function of review sentiment, product type and information quality. *Comput Hum Behav* 54:547–554
- Daud A, Ahmad M et al (2015) Using machine learning techniques for rising star prediction in co-author network. *Scientometrics* 102(2):1687–1711
- Duan W, Gu B et al (2008) The dynamics of online word-of-mouth and product sales—an empirical investigation of the movie industry. *J Retail* 84(2):233–242
- Forman C, Ghose A et al (2008) Examining the relationship between reviews and sales: the role of reviewer identity disclosure in electronic markets. *Inf Syst Res* 19(3):291–313
- Gao B, Hu N et al (2017) Follow the herd or be myself? An analysis of consistency in behavior of reviewers and helpfulness of their reviews. *Decis Support Syst* 95:1–11
- Ghose A, Ipeirotis PG (2011) Estimating the helpfulness and economic impact of product reviews: mining text and reviewer characteristics. *IEEE Trans Knowl Data Eng* 23(10):1498–1512
- Hong Y, Lu J et al (2012) What reviews are satisfactory: novel features for automatic helpfulness voting. In: *Proceedings of the 35th international ACM SIGIR conference on research and development in information retrieval, ACM*
- Hu Y-H, Chen K (2016) Predicting hotel review helpfulness: the impact of review visibility, and interaction between hotel stars and review ratings. *Int J Inf Manag* 36(6):929–944
- Huang AH, Yen DC (2013) Predicting the helpfulness of online reviews—a replication. *Int J Hum-Comput Interact* 29(2):129–138
- Huang AH, Chen K et al (2015) A study of factors that contribute to online review helpfulness. *Comput Hum Behav* 48:17–27
- Karimi S, Wang F (2016) Factors affecting online review helpfulness: review and reviewer components. *Rediscovering the essentiality of marketing*. Springer, Berlin, p 273
- Kim S-M, Pantel P et al (2006) Automatically assessing review helpfulness. In: *Proceedings of the 2006 conference on empirical methods in natural language processing, Association for Computational Linguistics*
- Korfiatis N, García-Bariocanal E et al (2012) Evaluating content quality and helpfulness of online product reviews: the interplay of review helpfulness vs. review content. *Electron Commer Res Appl* 11(3):205–217
- Krishnamoorthy S (2015) Linguistic features for review helpfulness prediction. *Expert Syst Appl* 42(7):3751–3759
- Lee S, Choeh JY (2014) Predicting the helpfulness of online reviews using multilayer perceptron neural networks. *Expert Syst Appl* 41(6):3041–3046
- Lee S, Choeh JY (2016) The determinants of helpfulness of online reviews. *Behav Inf Technol* 35(10):1–11
- Li X, Hitt LM (2010) Price effects in online product reviews: an analytical model and empirical analysis. *MIS Q* 34(4):809–831
- Liu J, Cao Y, et al (2007) Low-quality product review detection in opinion summarization. In: *EMNLP-CoNLL*
- Liu Y, Huang X et al (2008) Modeling and predicting the helpfulness of online reviews. In: *2008 8th IEEE international conference on data mining, IEEE*
- Liu Z, Park S (2015) What makes a useful online review? Implication for travel product websites. *Tour Manag* 47:140–151
- Lozano-Diez A, Zazo R et al (2017) An analysis of the influence of deep neural network (DNN) topology in bottleneck feature based language recognition. *PLoS ONE* 12(8):e0182580
- Malik M, Hussain A (2017) Helpfulness of product reviews as a function of discrete positive and negative emotions. *Comput Hum Behav* 73:290–302
- Malik M, Hussain A (2018) An analysis of review content and reviewer variables that contribute to review helpfulness. *Inf Process Manag* 54(1):88–104
- Malik M, Iqbal K (2018) Review helpfulness as a function of Linguistic Indicators. *Int J Comput Sci Netw Secur* 18(1):234–240
- Mudambi SM, Schuff D (2010) What makes a helpful review? A study of customer reviews on Amazon.com. *MIS Q* 34(1):185–200
- Nelson P (1970) Information and consumer behavior. *J Polit Econ* 78(2):311–329
- Ngo-Ye TL, Sinha AP (2014) The influence of reviewer engagement characteristics on online review helpfulness: a text regression model. *Decis Support Syst* 61:47–58
- Ngo-Ye TL, Sinha AP et al (2017) Predicting the helpfulness of online reviews using a scripts-enriched text regression model. *Expert Syst Appl* 71:98–110
- Otterbacher J (2009) ‘Helpfulness’ in online communities: a measure of message quality. In: *Proceedings of the SIGCHI conference on human factors in computing systems, ACM*

- Pan Y, Zhang JQ (2011) Born unequal: a study of the helpfulness of user-generated product reviews. *J Retail* 87(4):598–612
- Pang B, Lee L et al (2002) Thumbs up?: Sentiment classification using machine learning techniques. In: Proceedings of the ACL-02 conference on Empirical methods in natural language processing, vol 10, Association for Computational Linguistics
- Qazi A, Syed KBS et al (2016) A concept-level approach to the analysis of online review helpfulness. *Comput Hum Behav* 58:75–81
- Quaschnig S, Pandelaere M et al (2015) When consistency matters: the effect of valence consistency on review helpfulness. *J Comput-Mediat Commun* 20(2):136–152
- Samha AK, Li Y et al (2014) Aspect-based opinion extraction from customer reviews. arXiv preprint [arXiv:1404.1982](https://arxiv.org/abs/1404.1982)
- Singh JP, Irani S et al (2017) Predicting the “helpfulness” of online consumer reviews. *J Bus Res* 70:346–355
- Stieglitz S, Dang-Xuan L (2013) Emotions and information diffusion in social media—sentiment of microblogs and sharing behavior. *J Manag Inf Syst* 29(4):217–248
- Thelwall M, Buckley K et al (2010) Sentiment strength detection in short informal text. *J Am Soc Inf Sci Technol* 61(12):2544–2558
- Tsur O, Rappoport A (2009) RevRank: a fully unsupervised algorithm for selecting the most helpful book reviews. In: ICWSM
- Ullah R, Zeb A et al (2015) The impact of emotions on the helpfulness of movie reviews. *J Appl Res Technol* 13(3):359–363
- Ullah R, Amblee N et al (2016) From valence to emotions: exploring the distribution of emotions in online product reviews. *Decis Support Syst* 81:41–53
- Xiang C, Sun Y (2016) Research on the evaluation index system of online reviews helpfulness. In: 2016 13th international conference on service systems and service management (ICSSSM), IEEE
- Xu R, Xia Y et al (2008) Opinion annotation in on-line Chinese product reviews. In: LREC
- Yan X, Wang J et al (2015) Customer revisit intention to restaurants: evidence from online reviews. *Inf Syst Front* 17(3):645–657
- Yang S-B, Shin S-H et al (2016) Exploring the comparative importance of online hotel reviews’ heuristic attributes in review helpfulness: a conjoint analysis approach. *J Travel Tour Mark* 34(7):1–23
- Zhang J, Piramuthu S (2016) Product recommendation with latent review topics. *Inf Syst Front* 20(3):1–9
- Zhang Z, Varadarajan B (2006) Utility scoring of product reviews. In: Proceedings of the 15th ACM international conference on Information and knowledge management, ACM
- Zhang Z, Wei Q et al (2014) Estimating online review helpfulness with probabilistic distribution and confidence. *Foundations and applications of intelligent systems*. Springer, Berlin, pp 411–420
- Zhou S, Guo B (2017) The order effect on online review helpfulness: a social influence perspective. *Decis Support Syst* 93:77–87

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.