Classification of E-commerce online auctions and prediction of seller reputation

[1*]B. Valarmathi
Department of Software and Systems Engineering
School of Information Technology and Engineering
Vellore Institute of Technology
Vellore, India
valargovindan@gmail.com

[2]A Kavish pandit
Department of Information Technology
School of Information Technology and Engineering
Vellore Institute of Technology
Vellore, India
kavish.pandit2017@vitstudent.ac.in

[3]Pyla charan kumar
Department of Information Technology
School of Information Technology and Engineering
Vellore Institute of Technology
Vellore, India
pyla.charankumar2017@vitstudent.ac.in

[4]T. Chellatamilan
Department of Information and Engineering
Vellore Institute of Technology
Vellore, India
chellatamilan@gmail.com

[5]Abid Yahya
Professor,

Electrical, Computer & Telecommunication Engineering Department
College of Engineering & Technology
Botswana International University of Science & Technology
Palapye, Botswana
yahyabid@gmail.com

Abstract— Seller reputation plays a vital role in the E commerce market. The significant profit and analysis of the overall giant venture which provides both observable and heterogeneous reputation for the online community sellers. Greater validation and performance outcomes to be provided in the optimal algorithm preferred. The methodology that makes the outcome of a multinominal naïve bayes highly efficient, accurate and also enhances a better understanding in reference to E-commerce market and reputation by using singular value decomposition (SVD) algorithm. Online experiments and empirical studies on the data from an online auction market to show that reputation can generate cut-offs for sellers and to generate the detailed analysis of the reputation systems giving a proper boost to the online e-commerce community.

Keywords— Auction, Multinomial Naive Bayes, SVD, E-commerce, Reputation system.

I. INTRODUCTION

E-commerce website system heavily relies on the two major factor which includes trust and reputation. Reputation enhances not only the seller end but it generates the customer perspective regarding the online community base. Trust is used by an individual to assess the performance of another individual by his actions. Thus it becomes necessary to incentivize the sellers to be honest. Reputed online commerce websites such as Alibaba, Amazon, eBay generate tremendous economic value with huge transaction volume.

E-commerce systems serve as online shopping markets, where buyers can purchase products from online stores, each of which is operated by a seller. Further, it becomes very important to enhance the user interaction for the success of the E-commerce .Including a multi-dimensional approach in order to obtain a fine grained trusted computation. It is critical for the buyers to enhance the summarization and opinion extraction. It is critical for the buyers to enhance the summarization and opinion extraction.

Digging further into the concept of the performance outcomes to be provided, there is a need to apply the optimal algorithm .Hence there is a need to conduct experiments and figure out the appropriate classification techniques on the global dataset and generate the required outcomes.

Reputation models on the other hand provides the reputation and trust scores for further implementation of the trust metric. The paper involves the matrix factorization using the singular value decomposition for splitting the matrix into the set of factors as well.

Collaborative filtering and multi-dimensional trust models comes under the known application of the SVD algorithm and it considers the correlation between different data item sets.

Thus this article focuses on the best framework for the e-commerce reputation and defining the optimal classification techniques on the auction dataset with number of auction days as the class labels.

II. LITERATURE SURVEY

Techniques used for existing systems are mentioned below:

A. Ecommerce online auctions
B. Seller Reputation
C. Optimization Algorithm
D. Matrix Factorization technique applications.

A. Ecommerece online auctions

Hong xie et. al. [1] proposed the online Auction System is a practice followed by the Ecommerce sites which takes place between auctioneers and bidders. This enhances the growth from seller's side in terms of popularity, brand value with low term profits. All the transactions made during the auction are categorized into a dataset for further analysis. Different machine learning techniques are then applied on categorized data for performance outcomes. Both Auctioneers and bidders get simultaneous benefits as a seller reputation enhances with bidder requirement satisfaction. Ramanthan Guha et al. [9] proposed a model for a web of trust, a user can prepare and view of some other user without prior communication.

B. Seller Reputation

Daniel Houser and John Wooders [2] suggested the various sellers involved in the online auction to enhance their respective fields of reputation. Each seller have their own reputation, but in hierarchical manner when grouped together. This hierarchical manner can be visualized and analyzed by a technique named "Matrix Factorization technique applications". Concept related to linear algebra (Singular value decomposition) to be used for indexing and prioritizing the reputation scores of sellers. Vector representation is used to represent the user and product interactions. For effective prediction of ratings in collaborative filtering, standard SVD is used. [7] fortune ranking data are retrieved from fortune 500. Ginger Zhe Jin et al. [10] mentioned that the online seller reputation is active for finding worthy sellers. In any case, restrictive on completed auctions, trustworthy venders don't give better quality. Stuart Landon et. al. [11] idea was taken which initialized the concept of quality expectations in ecommerece auction markets through the sellers reputation and prices.

C. Optimaztion Algorithm

Machine learning has enjoyed tremendous success and this success can be attributed to the data driven philosophy that underpins it. Machine learning algorithms are widely used in numerous datasets but there is a constant need to provide the accurate results. This brings the concept of optimization and relies on some selective algorithm.

Our paper deals with numerous attributes but we selectively focus on the auction type to classify the data. We arrive at the conclusion by applying the Multinomial naïve bayes classification and get the enhanced seller results. Thus the seller reputation is analyzed by classifying the data and comparing it with the previous bidder's rates. This further improves the ranking and increases the clear visibility for the preferred outputs. And making use of the basic ranking criteria also it arrives on the ecommerce reputation as well

D. Matrix Factorization technique applications.

Sulin Ba et al. [4] used the singular value decomposition (SVD). It is an one of the matrix factorization techniques and a unsupervised algorithm that to be applied on seller reputations to arrange the positions of popularity in a hierarchical manner. The seller at top of the hierarchical order is more reputed and recognized as the highly socialized seller among all the online community sellers.

The working mechanism of the Singular Value Decomposition follow as: Reputation scores or values are considered as the numerical entries and represented in a matrix. This matrix is further factorized into 3 matrices. U, Sigma ,V transpose are three matrices that are splitted matrices obtained by the singular value decomposition. Values in the column of the matrix U, gives the preferences or hierarchies over all the respective reputation scores. This is the concept which flows in the background to visualize and recognize the seller's popularity. John Gittins et al. [8] proposed the bandit process and described the dynamic allocation indices which was used as a reference for allocation of dynamic numeric indices for obtaining the rank and reputation.

III. BACKGROUND KNOWLEDGE

Enhancing the reputation systems by providing a pertinent solution and removing the fraudulent behaviour gives the E-commerce website a secure online base. This gives a better narrative to the upcoming research for our paper and hence enhance the decision making for the bidder. For instance a bidder will place a bid only if his/her value for the item would exceed the higher bid at the time and he/she first would consider viewing the auction. Also the count of bidders is mostly larger than the overall bidding done.

Moreover the best measure for the count of bidders might be the number of users who can click on the link to the auction's main page. Thus the link not only displays the current price, it would be focusing in undercounting the number of bidders. The best measure of the number of bidders might be the number of users who click on the link to the auction's main page. The link provides the seller's reputation which defines in the way but provides it in the count of bidders affected by the seller's reputation.

Peter Auer et. al. [3] suggested the seller may choose an auction length of 1, 3, 5, 7 days when selecting an item. For example, consider an online reputed sire eBay which includes the largest concurrent auctions are conducted at one go among 40 million buyers by the auctioneers and seller behind the scenes expecting the reputation.

If eBay is taken as reference for the above case, it doesn't provide any warranty for auctions, it acts as a listing service. Buyers in this case, find that there is some risk correlated with it. Even fraudulent transactions are expected.

We as a team performed a task of finding out the optimal algorithm for the given dataset and to feature our results by finding out the relation between auction and the seller reputation. India online ecommerce base is plagued by several problems, and they are easily facilitated by the different fraudulent behaviours alongside.

Certain situations, there may be one way to find out an compare the raters reports to find out the referred peer reports and thus generate reward agreement. However, if the rewards are made part of the process, there may be certain chances of the danger. So, if there is a certain outcome which is most likely to occur, such as the positive experience with a seller at eBay who has a stellar feedback past experience, then there is a rater who has a bad experience which also expects that the next rater is likely to have a good experience.

Chrysanthos Dellarocas [5] suggested many online systems, however, raters seem to be quite motivated by prestige or privileges within the system. To the best of our knowledge, this was the first time that has corrected for this difference. And no previous study has corrected for this difference. This is because of measure mentor or in reputation variables, previous estimates of seller reputation effects are highly inconsistent and biased as well.

[6] According to the eBay website and it's analysed research, the auction contract develops the relation between the seller and the bidder which is both binding and have the utmost importance.

Henceforth, according to eBay without paying the item bid on Neither eBay nor the other main consumer-to-consumer auction sites provide away for bidders to post a depositor payment penalty in the event of default. We did emphasize on the seller reputation (but not bidder reputation) which significantly determines the auction prices.

Reputation recognized as a significant evolutionary part in online community auctions positive feedback from a user or buyer enhances the reputation of a seller community. Seller looks forward for emphasizing on this areas to build up his side from user perspective. This make sellers to take incentives for greater performances in their respective auctions conducted.

A deep study of how the process of reputation is going to build incentives for better performance in contracts is a key area for future researchers.

IV. EXPERIMENTS AND SYSTEM DEVELOPED

Fig. 1 emphasizes the various kinds of exploration done by us to obtain a optimized classification techniques among the all others applied up to know. This flow has a simultaneous techniques applied to focus on enhancing the seller reputation and ranking in an optimized fashion.

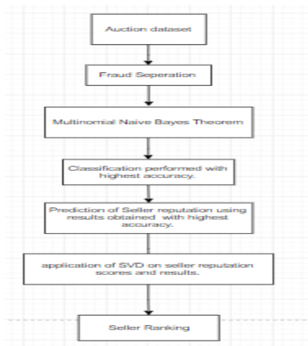

Fig. 1. The flow chart of proposed work

Auction data set is the most compatible for analysis of bid and bidder statistics which play a major role for detecting the seller reputation through online seller community auctions. 'Auction type' attribute from the dataset is taken as a class label to detect Auction day type.

Further the ranking generated is formulated through the SV algorithm and thus helps in clear division of the seller reputation in the online base. The classification performed also requires the highest accuracy. We can explain the concept by starting with the seller base which gets the direct benefit of the online auctions.

E-commerce websites have trusted sellers across the platforms who are an active member of the online community. Their reputation on the other hand increases by the help of the online auctions and this eventually helps in the transactional volume on the website as well. Further after the increase in the seller reputation we arrive to the increase in the popularity of the online sites as well. Thus not only auctions help in increasing the reputation of the seller but it generates the required popularity for the e-commerce website a well. Auction data thus plays a 'major' role in the whole scenario and provides a efficient online environment for the sellers and the bidder as a whole.

Multinomial naive bayes theorem and methodology is applied on the online auction dataset. Table 1 shows results obtained from total number of instances [10681] for the accuracy of multinominal naïve bayes classifier.

TABLE 1: ACCURACY OF MULTINOMINAL NAÏVE BAYES CLASSIFIER

Description	Rows	Accuracy
Correctly Classified Instances	**10632**	**99.5412 %**
Incorrectly Classified Instances	49	0.4588 %

Through the performance outcomes of the classification technique applied, 10632 instances are correctly classified out of 10681 with an accuracy of **99.5412 %**. Naïve bayes multinomial theorem would be most efficient, preferable and optimized algorithm for classifying auction dataset and attributes. It would be recommended as a major task for analyzing online community auctions and result in best performance outcomes for further analysis of reputation and rankings of seller's community alongside.

V. DATASET DESCRIPTION

The dataset contains 10682 rows and it consists of 9 basic attributes as mention in the columns.

All the features in Auction dataset are described below.

a) Auction id: Each Auction has a unique identity represented by a identifier
b) Bid: the primary bid placed by a bidder in an auction.
c) Bid time: The time taken in days for a bid to get placed, from the start of the auction.
d) Bidder: Bidders eBay username.
e) Bidder rate: eBay provides a feedback option for bidders rating.
f) Open bid: It represents the open bid set by the respective seller
g) Price: The final price to which the item is sold. (Refers to the second highest bid+ a gradual increment.)
h) Item: auction item
i) Auction type: Representing the days of auction.

Fig..2 shows the graphical analysis between Bid value (Highest price) **a** buyer of a stock is willing to pay in x-axis and Bidder rate (Base currency) in y-axis

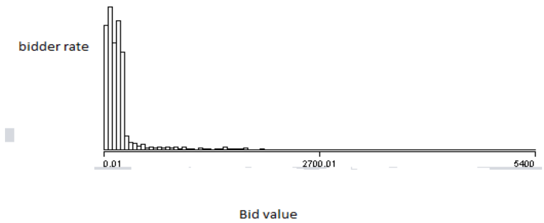

Fig. 2. Bid value versus Bidder rate

Fig. 3 shows the graphical analysis between Bidder time (Time taken for receiving the bid request and serving) in x- axis and Bidder rate (Base currency) in y-axis

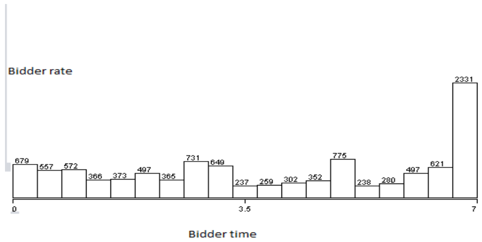

Fig. 3. Bidder time versus Bidder rate

Fig. 4 shows the graphical analysis between Items Available in Auction representing x-axis and y axis represents the count of bidder's participation in auction.

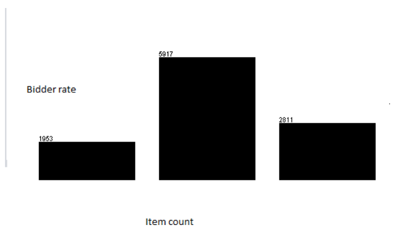

Fig. 4. Bidder count versus Item count

Fig. 5 shows the graphical Analysis of day type's (3 day, 5 day, 7 day) in x-axis and y-axis represents count of bidders participated in a particular day.

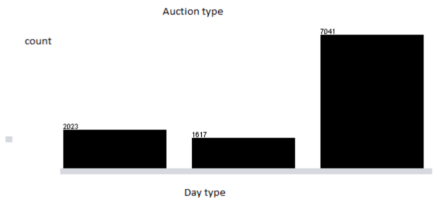

Fig. 5. Day type versus Count

V. PROPOSED WORK

Our research deals with the bidder seller relationship in the online world and shows the various possibilities that occur. This expands the further proposed work and relates it to the derived content as follows. Thus it includes various functionalities which are to be involved and mentioned below.

Steps involved in application of Multinomial Bayes Algorithm.

1. Read the auction dataset.
2. Formulate the mean and standard deviation of the predictor variable in different class of the auction dataset.
3. Repeat the following steps:

Formulate the probability of f, using the Gaussian density equation in each class. Until the probability of all predictor varaibles (f1,f2….fn) has been calculated and the likelihood for each class.

TABLE 2: COMPARISON OF DIFFERENT CLASSIFICATION ALGORITHMS

S.No.	Classification Technique	Accuracy obtained	Number of Instances classified
1.	Zero R	65.92%	7041
2.	Input Mapped Classifier	65.84%	7041
3.	Multinomial naïve Bayes Theorem	99.54%	10632

The different classification algorithms is applied on auction

dataset such as Zero R, Input Mapped Classifier and Multinomial naïve Bayes Theorem. Table 2 shows the comparison of the accuracy results generated from the code of corresponding algorithm applied. Our research paper thus gives the highest accuracy for the multinominal naïve bayes and generates it for the 10,000+ instances alongside.

VII. FUTURE WORK

Any further advances in present paper concepts can be encouraged and those advances are kept remained for future due to less sufficiency of time, which is therefore linked with actual data and a time consuming process. This involves in deeper analysis of present concepts. It builds a new curiosity and ideas of evolution for further advances of present concepts when deep analysis is performed among the presented concepts in this paper.

Results obtained through experiments in this paper are quite satisfactory. Even further study and application of techniques is required in order to understand the functionalities to be developed or improved.

VI. CONCLUSION

Our results provide a simple and a generic framework that gives a detailed analysis on the effect of reputation and auction data. Further the use of the optimized algorithm by applying the multinomial naïve bayes algorithm gets the highest preferred outcome .The implications are being put on the best possible way.

Apart from the classification algorithm this paper addresses matrix representation technique in the form of singular value decomposition. The generated results have an accuracy of about 99.54% .We built all the histograms on the optimal attributes and scanned with the feedback system which thus provides incentives for good performance.

On the other hand, our paper provides clear knowledge on the relationship between the auction type and the seller reputation. To build trust between any two people in the system with higher accuracy, we show the distrust rounding and other types of phenomenon which will further have significant effects on how trust is propagated as well.

VIII REFERENCES

[1] Hong xie and Richard T. B.MA, John c. s. Lui, "Enhancing Reputation via Price Discounts in E-Commerce Systems: A Data-Driven Approach", 2018.

[2] Daniel Houser and John Wooders, "Reputation in Auctions: Theory,and Evidence from eBay.2000. Journal of Economics & Management Strategy", Vol. 15, pp. 353–369, 2006.

[3] Peter Auer, Nicolo Cesa-Bianchi, and Paul Fischer, "Finite-time analysis of the multiarmed bandit problem", Journal of Machine Learning, Vol. 47, pp. 235–256, 2002.

[4] Sulin Ba and Paul A. Pavlou, "Evidence of the effect of trust building technology in electronic markets: Price premiums and buyer behaviour", vol.

26, pp. 243–268, 2002.

[5] Chrysanthos Dellarocas,. "Analyzing the economic efficiency of eBay-like online reputation reporting mechanisms", In Proc. of ACM EC, 2001.

[6] eBay. 1995. Homepage. Retrieved from http://www.ebay.com/.

[7] Fortune 500. 2015. Fortune Ranking. Retrieved from http://fortune.com/fortune500/.

[8] John Gittins, Kevin Glazebrook, and Richard Weber. 2011. Multi-armed Bandit Allocation Indices. John Wiley & Sons.

[9] Ramanthan Guha, Ravi Kumar, Prabhakar Raghavan, and Andrew Tomkins. 2004. Propagation of trust and distrust. In Proc. of WWW. 403–412.

[10] Ginger Zhe Jin and Andrew Kato, "Price, quality, and reputation: Evidence from an online field experiment" , The RAND Journal of Economics vol. 37, pp. 983–1005, 2006.

Stuart Landon and Constance E. Smith, “Quality expectations, reputation