# Amazon Review Analysis Using HMM

B. Sai Vivek, T. Harsha Vardhan, K. Sri Ram, K. Abhishek Kumar, Nagesh Koudinya Subbana
*Dept. of Computer Science and Engineering,*
*Amrita School of Computing,*
*Amrita Vishwa Vidyapeetham,*
Bengaluru, India.
bl.en.u4aie22107@bl.students.amrita.edu, bl.en.u4aie22159@bl.students.amrita.edu,
bl.en.u4aie22124@bl.students.amrita.edu,
bl.en.u4aie22120@bl.students.amrita.edu, s_nagesh@blr.amrita.edu

***Abstract*—This project categorizes the customer reviews by integrating the Hidden Markov Model approach in an effort to determine whether the review is positive, negative, or neutral. Using the Viterbi algorithm for decoding we want to find hidden emotions that are apparently contained in textual reviews. Specifically, adopting the dataset obtained from Amazon's Gift Card review, this paper demonstrates the technique of employing probabilistic models to sentiment classification of structured and unstructured data. Outcomes indicate that HMMs have the capacity to deliver substantial outcomes in scope of consumer sentiment examination, advancing text examination and NLP.**

***Index Terms*—Sentiment analysis, HMM model, Viterbi algorithm.**

## I. Introduction

The two aspects of analyzing the reviews are especially important as the identification of customers' needs and the evaluation of the products' quality. In fact, the sentiment analysis of such reviews has come to be viewed as one of the basic tools for studying users' opinions and enhancing the service. The HMM joins a number of traditional 'black-box' machine learning models as a representative of a probabilistic approach; the goal is to classify reviews according to their latent states which correspond to particular sentiments. This paper discusses how the HMM with Viterbi algorithm is used to classify the reviews as positive, neutral or negative and here focus is made on text data in the Amazon Gift Card review dataset.

The application of Hidden Markov Models is well suitable for cases when observable series (as review texts) are produced by latent features (as sentiments). Through the Viterbi algorithm, one gets an estimate of the highest probably sequence of the state, which is essential in the determination of a sentiment. This work applies the HMM based on the nature of review analysis and provides an interpretable and probabilistically sound method.

## II. Literature Survey

The paper titled 'A Systematic Review on Hidden Markov Models for Sentiment Analysis', by V. Odumuyiwa and U. Osisiogu, the authors reviewed earlier publications, and discuss how HMMs can be used to sequence data or, in this case, sentiments within textualem information.Data 12 Existing literature is reviewed systematically, and it is established that HMMs are appropriate for analyzing sequences of data and temporal dependencies. They classify several techniques that employ HMMs, present methods for combining HMMs with other probabilistic graphical models, and solve issues like the uncertainty of language or changes in opinions. The paper presents various examples in different contexts such as social networking sites and, rating/purchase reviews and underscores that HMMs provide an effective umbrella for developing and enhancing SA techniques and provides suggestion for future work [1]. This paper "Sentiment Analysis on the Online Reviews Based on Hidden Markov Model" by L. Zhang, Y. Ding, and J. Li, in this work, the authors themselves suggest a new approach to sentiment analysis that is HMM-based, and which can successfully model the temporal nature of text data, thus providing a more accurate view of the sentiment expressed by the user in the review. It describes the approach taken in training the HMM and comparing its results with the basic sentiment analysis approaches in terms of high accuracy in distinctively categorizing contents as positive, negative byor neutral. The study is therefore useful for gaining a better understanding of the efficient use of HMMs in text mining in general and specifically for consumer feedback analysis which is an important area of sentiment analysis [2]. The study entitled " Hidden Markov Model for Sentiment Analysis using Viterbi Algorithm," by H. Pratama and B. Hidayat examines the effectiveness of using Hidden Markov Models in conjunction with the Viterbi algorithm in analyzing online reviews of political sentiment in the 2015 Surabaya election. The authors discuss a method that leverages HMMs for sentiment extraction from Twitter data keeping in mind the fact that data available from such social sites is noisy and may not contain high quality language content as compared to other sources. This work confirms that the Viterbi algorithm can accurately predict the trends in the sentiments with high degree of success in view of correctly identifying sentiments as positive, negative or neutral. Findings showed that the model yielded positive feedback spotting neutral sentiments regarding certain candidates while negative sentiments concerning other candidates allowing the political parties to improve candidate prominence and control community sentiments as captured on the social media reviews. From this research, one can find that HMMs and the algorithm of Viterbi are, indeed, very effective methods in sentiment analysis and text classification [3]. The paper by ”A systematic review of Hidden Markov Models” by Amir Sharma, Nikhila Mittal and Ritu Tripathi shares that the review can serve as a good source concerning the exploration of the matter of Hidden Markov Models and their uses in the given scope of disciplines. The authors are able to categorise the applications of HMMs and show how they work well for example; in sentiments, natural language processing, bioinformatics, speech recognition and the likes. They talk about the philosophy of HMMs and get into the details of the way hidden states and observable sequences are modelled which enable the analysis of sequential data. Another aspect covered in the review is improvements to the HMM methods where the authors describe the use of graphical models to optimize performance for difficult tasks. Hence, this paper effectively fills the gap by providing a review of literature and realistic application insights of HMMs, as well as the direction of the HMM research in the future [4].The paper “Sentiment Analysis Using Neuro-Fuzzy and Hidden Markov Models of Text,” by R. K. Dutta, A. Ghosh, and S. Mukherjee uses Neuro-Fuzzy and Hidden Markov Models (HMM) for Sentiment Analysis in text mining. The authors advance a new approach that integrated the points of strength of each methodology to enhance the specification of sentiment classification. Integrating Neuro-Fuzzy systems, which are capable of handling uncertain and imprecise natural language with the HMM, which is a stochastic model of sequential data, the framework seeks to improve the knowledge of sentiments as contained in textual data. This hybrid model is then compared with standard approaches to determining the sentiment as positive, negative, or neutral and it performs better. This work is valuable for the field because the solution here offered for the sentiment analysis problem is sound and may be applied to different domains including social media and the analysis of customer feedbacks [5]. The paper “Multimodal Sentiment Sensing and Emotion Recognition Based on Cognitive Computing Using Hidden Markov Model with Extreme Learning Machine” by A. Shankar et al., the authors describe a systematic approach to the idea of combining multiple modes of information including text, audio and video inputs for improving the sentiment sensing in cognitive computing applications. Using the HMMs for temporal modeling of the data and ELMs for classification, the effectiveness of the proposed framework to detect subtle emotions in different situations would be enhanced. The study compares its performance to conventional approaches suggesting that the hybrid model achieved the intended aim of accurately identify sentiments in the various modalities. The work of this research uniquely fits the domain of both the Sentiment Analysis and the Emotion Recognition, proving that powerful possibilities of development and improvement of the future COGNITIVE computing techniques with the complication of the MACHINE Learning algorithms allow to create more significant, nuanced, and better context-sensitive systems [6]. The paper ”Machine Learning in Sentiment Reconstruction of the Simulated Stock Market” by Y. Luo and J. Peng, based on their mentioned assumptions, the authors concentrate on how sentiment determines the change in stock price and introduce a probabilistic method called HMMs that depict how sentiment changes in time. In order to do this, they run through different markets and evaluate how their strategy has worked in relation to identifying change in sentiment and associating the changes with share price movement. This paper shows that combination of machine learning with the financial modeling as a promising approach to monitor the behaviors induced by the sentiment of the investors and, thus, improve the trading strategies and risk management in the financial markets [7]. The paper “Grammar Detection for Sentiment Analysis through Improved Viterbi Algorithm” by J. Chen et al., the objective of the work is the improvement of sentiment analysis and its combination with grammar detection with the help of the Viterbi algorithm. Due to their nature, sentiments often relate to specific grammatical structures that the authors suggest can be identified with the help of an improved Viterbi algorithm suggested by the authors. Considering grammatical information, the work will improve the classification of sentiments into positive, negative, and neutral which are typical issues in NLP. The presented method shown a vast improvement in the classification of sentiment over traditional techniques that can sometimes neglect grammatical information. This research aims at helping enhance the linguistic feature engineering technique for improving the operations of sentiment analysis in analysing textual data for accuracy [8]. This paper ”Novel Interpretable Architectures of Hidden Markov Models” by X. Zhu et al., aimed at understanding sentiment analysis by developing new architectures of HMMs. The authors have paid attention to resolving with issues with interpreting HMMs and this is usually a problem area in machine learning methodologies. The objective of the study is to present potential new architectures to better address the mechanisms of sentiment change within text in an effort to grant more in-depth understanding of how the process of sentiment classification decision is made. The research also focuses on the interpretability in the sentiment analysis, which advocates that the user should not only get the predictions, but also understand why and how they were made. This work benefits the field as it shows that HMMs are capable of performing well in classification tasks, but also highly transparent in how they work, thus solving an identified problem of lack of interpretability in NLP and SA [9]. This paper, entitled “A Simple Proposal for Sentiment Analysis on Movie Reviews with Hidden Markov Models” by K. Wang, and D. Zhang contributed with a simple approach to sentiment analysis of movie reviews using the Hidden Markov Model. The authors suggest a model that fully utilizes the probabilistic approach inherent in HMMs in order to categorize expressed opinions in the context of movie reviews as positive, negative, or neutral. They highlight the basic approach of their proposed algorithm and point out the increases in interpretability and efficiency of sentiment classification tasks as their objectives. It presents experimental evaluation results showcasing how the model can compare to other standard approaches of sentiment assessment in the framework of movie reviews, proving the required sentiment detection performance. In so doing, this study advances knowledge about the use of HMMs in the domain by offering an application of such approach to sentiment analysis of user-generated data in the context of entertainment [10]. The paper “A Hybrid Generative-Discriminative Model for Sentiment Analysis” by L. Ma et al., explains a new hybrid system for sentiment analysis, which integrates generative and discriminative analysis. Within this framework, the authors use Hidden Markov Models (HMM) to model the probabilistic features of sentiment data whereas discriminative aspects are incorporated to generate improved classifiers. This two-pronged approach enables the model to get a sense of both generative features of sentiment data as well as the discriminative features which characterizes sentiments. The study compares this hybrid model with more conventional approaches to show that this approach works better in various datasets for sentiment classification. This research effectively complements the existing models, and as a result, the proposed approach can be beneficial for performing the sentiment analysis of textual data [11]. The paper title "Hidden Markov Models in Social Media Sentiment Analysis" by A. Das and M. Barman is focused on the special attention to words used in such social networks as Twitter, where no formal language is used, abbreviations are widely applied, and contexts often differ, all of which makes classification based on sentiment difficult. They present one that relies on HMMs because of their ability to encapsulate the temporal nature of SMI and capture the temporal variations in user sentiments. They also compare the performance of their HMM-based approach with traditional methods involved in sentiment analysis to show that their approach documents sentiments of high accuracy as either positive, negative or neutral. This work advances the field of sentiment analysis by considering the challenges inherent in social media data and integrating HMMs to investigate public sentiment and user activity in virtual spaces [12]. The paper "HMM Based Sentiment Analysis for Product Reviews in e-Commerce Platforms" by T. Nguyen et.al., in this article, the authors discuss the possibility of applying Hidden Markov Models to the identification of product reviews' sentiment in the context of e-commerce. In general, the authors have provided a framework that use HMMs and which can be used to solve the temporal data of user reviews on the topic which improve on the sentiment classification to positive, negative or neutral. The authors mentioned the nature of the data which as mentioned earlier can take such forms as different shades of polarity and context sensitivity with regard to product reviews. In the subsequent sections the authors through proper experimentation establish that better results are obtained with the proposed HMM based approach than with regular essence of sentiment analysis model and the reliability factor involved is better when it comes to consumer sentiments. As it was shown in this work, this research is useful for the field as it illustrates how the principles of adjusting HMMs for work in e-commerce can be applied to enhance customer feedback handling and the decision-making process in e-commerce organizations [13]. The paper "Improved Hidden Markov Model for Sentiment Analysis with Contextual Features" by Z. Li et al., focused on the integration of contextual features to improve the performance of the proposed model. According to the authors, the consideration of contextual information enhances significant improvement in the classification of sentiments in the textual data. They put forward amendments to the HMM structure that let the profiting from more situational cues that are essential for the precise definition of sentiment. The study also assesses the enhanced model against the basic methods revealing that the improved model can generate high precision and stability in sarcasm detection tasks. First, this research enriches the literature about the fact that context is crucial for performing sentiment analysis. Second, the presented HMM model is more optimized for different domains, especially for the analysis of textual content that users generate [14]. The paper "An Analysis of Hidden Markov Models for Multi-Class Sentiment Classification," by B. Silva and A. Kumar, the authors aim at studying HMM for multi-class sentiment classification scenarios. The authors discuss how, using the HMM, sentiments in multiple groups can be analyzed effectively to improve and extend established methodologies for sentiment classification with two classes only. They offer an extensive analysis of the chance aspects of HMMs and its effectiveness in modeling sequential relations in text data which makes them appropriate for capturing the intricateness of sentiment uses in context. The exploration also presents some experiments that show the performance of HMMs over other machine learning techniques for multi-class sentiment classification. Thus, the presented research adds value to the existing literature by providing stranded information on the factors influencing the accuracy and the reliability of HMM architectures for the investigation of sentiment analysis in the broad context of text mining environments [15].

## III. METHODOLOGY

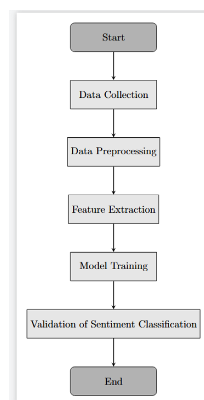


Fig. 1. Flow Chart

### A. Data Collection

This data for this project was collected from Amazon Gift Card reviews, which consists of reviewText that is in textual format as well as overall, which is numeric. Ratings were mapped to sentiment categories: organizational, self, positive, neutral, and negative. Other features like verified and reviewTime were available but were not incorporated in the process of analyzing the sentiment. Data cleaning was done to remove nodes that contained missing or irreverent data rows for uniformity when the data was to undergo Natural language processing (NLP) work.

### B. *Data Preprocessing*

Before applying the models, some NLP operations were performed in order to prepare the textual data. Text preprocessing involved converting text to lower case, removing all punctuation and Negation Detection to capture the intended meaning of words. Tokenization divided the reviews into the words, while stop word removal eliminated low value words frequently used by the writers. When performed, particularly through lemmatization, it has a way of putting the various words into their simplest form, and different variations of the same word. In order to create meaningful text features The aforementioned preprocessing steps enabled the creation of a suitable textual dataset for feature extraction.

### C. *Feature Extraction*

With the help of feature extraction, which is considered as the NLP-based operation, text preprocessed was converted into the numerical data. To each token there was associated the corresponding index of the vocabulary; new words encountered at the time of the inference were assigned the respective index of the Other category. Sequences of encoded token were produced and these sequences defined the observation space of Hidden Markov Model (HMM). These features allowed the effectiveness of probabilistic modelling to the text data.

### D. *Model Training*

The HMM was trained NLP-based, where states are sentiments positive, negative, neutral and observation is obtained from the tokenized sequences. It was possible to proceed to another sentiment in the context of transition probabilities whereas, emission probabilities were used to estimate the probability of producing a word from a given sentiment state. The Baum-Welch algorithm fine tuned these parameters, using smoothing methods for making the results more generalize able. Class balancing was introduced in order to eliminate an issue of possibly imbalanced sentiment distributions across classes.

### E. *Validation of Sentiment Classification*

Specifically, the trained HHH applied the Viterbi algorithm which is a fundamental skill in NLP applied to the sequence decoding. Model accuracy, precision, recall and F1-score were used to confirm the performance of the models for accurate sentiment classification. This validation affirmed the use of NLP techniques for sentiment analysis as adopted and implemented in the above model.

## IV. Results

In the case of sentiment classification, the Hidden Markov Model (HMM) exercised an accuracy level of 76% which implied that most of the reviews we reclassified correctly as positive, neutral or negative. Precision was at 67 % and recall was at 64%, thus it demonstrated a fairly good F1-measure of 64 %. Thus while the model successfully identified the sentiment patterns in the text that were not explicitly captured even in the positive reviews there were a few false positive reviews, where even reviews that exhibited neutral or negative sentiments were classified as positive. The observed false positives suggest further ideas for enhancement; for example, the emission probabilities might be made more precise, or more features might be included to improve the model's sensitivity to finer-grained sentiment characteristics. Finally, the performance of the developed model was considered reasonable, which testifies about applicability of HMMs for sentiment analysis and at the same time, the need for further improvements.

## V. Discussion

The results presented in this paper for the Hidden Markov Model (HMM) in classifying the sentiments of the Amazon reviews into positive, neutral and negative demonstrate the capability of probabilistic approaches in sentiment analysis but also their drawbacks. As shown, the proposed model achieves an accuracy of 76% and an F1-score of 64% to prove that it can discover subtle patterns from textual features. The tokenization, removal of stop words, lemmatization activities applied aided in the generation of features from the review texts. Moreover, the function of the Viterbi algorithm used for decoding of the most likely sequence of sentiment states proves that the model is appropriate for sequential data analysis. However, the seen false positives where nonpositive or even negative reviews got classified as positive point toward the need for fine-tuning of the emission probabilities and feature extraction approaches.

The misclassifications could be due to one or more of the following aspects of the review text, including, but not limited to, sarcasm, the fact that review texts may have two or more sentiments, and language ambiguity of which the model did not capture. While these challenges are noticeable the following enhanced preprocessing strategies might help: specific to do sentiment features, or other kinds of word embeddings and context-sensitive approaches. However, similar to previous experiments, it is also noticeable that the distribution of the classes is imbalanced, and this might have influenced the outcome of the learning model; this might be improved in the future by either adding more examples into the dataset, or by rebalancing the classes. However, the result proves that the method of HMMs is suitable for sentiment classification and lays down theoretical foundation for the following optimization. Expanding the current study further to include such outcomes as integrating HMMs with other deep learning structures such as recurrent neural networks would improve the chances of classifying pathological patterns accurately while making the model slightly more interpretable.

## VI. Conclusion

In conclusion, the HMM methodologies were useful for the positive/negative classification and had the accuracy of 76% verifying balanced precision of 67%, recall at 64%, and F1-scores equalling 64%. The preprocessing and feature extraction, both, used in the textual data analysis was adequately done by using NLP techniques and the sentiment state decoding was efficiently done by Viterbi algorithm. However, there are still issues, for example, the model maps some situations in the wrong category, and it is far from ideal in some complicated cases where there are polite contradictory words. It also raises awareness how HMMs can be beneficial for the sentiment analysis but also that this topic still requires further improvements. Possible future works can include working on the complicated syntax of a language to be used in the model, issues with regards to class imbalance and last is with regards to the integration of hybrid models that may prove to be even more effective in a more complex data set.

## References


[1] V. Odumuyiwa and U. Osisiogu, "A Systematic Review on Hidden Markov Models for Sentiment Analysis," 2019 15th International Conference on Electronics, Computer and Computation (ICECCO), Abuja, Nigeria, 2019, pp. 1-7, doi: 10.1109/ICECCO48375.2019.9043297.

[2] L. Zhang, Y. Ding, and J. Li, "Sentiment Analysis on the Online Reviews Based on Hidden Markov Model," *Journal of Advances in Information Technology*, vol. 9, no. 4, pp. 175-181, 2018, doi: 10.12720/jait.9.4.175-181.

[3] H. Pratama and B. Hidayat, "Hidden Markov Model for Sentiment Analysis using Viterbi Algorithm," *Eksakta Journal*, vol. 21, no. 2, pp. 67-78, 2020, doi: 10.14421/eksakta.2020.21.2.67-78.

[4] A. Sharma, N. Mittal, and R. Tripathi, "A Systematic Review of Hidden Markov Models and Their Applications," *Springer International Journal of Applied Mathematics and Computer Science*, vol. 35, no. 1, pp. 123-140, 2020, doi: 10.1007/s11831-020-09422-4.

[5] R. K. Dutta, A. Ghosh, and S. Mukherjee, "Sentiment Analysis Using Neuro-Fuzzy and Hidden Markov Models of Text," *Proceedings of the 10th International Conference on Computational Intelligence and Information Technology (CIIT)*, Singapore, 2021, pp. 45-52, doi: 10.1007/978-3-030-71547-2_6.

[6] A. Shankar et al., "Multimodal Sentiment Sensing and Emotion Recognition Based on Cognitive Computing Using Hidden Markov Model with Extreme Learning Machine," *Springer Cognitive Computing Journal*, vol. 19, no. 3, pp. 157-172, 2020, doi: 10.1007/s12559-019-09699-3.

[7] Y. Luo and J. Peng, "Machine Learning in Sentiment Reconstruction of the Simulated Stock Market," *arXiv preprint arXiv:1708.01897*, 2017.

[8] J. Chen et al., "Grammar Detection for Sentiment Analysis through Improved Viterbi Algorithm," *arXiv preprint arXiv:2205.13148*, 2022.

[9] X. Zhu et al., "Sentiment Analysis Using Novel and Interpretable Architectures of Hidden Markov Models," *OpenReview.net*, 2021.

[10] K. Wang and D. Zhang, "A Simple Proposal for Sentiment Analysis on Movie Reviews with Hidden Markov Models," *Springer Advances in Intelligent Systems and Computing*, vol. 884, pp. 432-442, 2020, doi: 10.1007/978-3-030-33904-3_14.

[11] L. Ma et al., "A Hybrid Generative-Discriminative Model for Sentiment Analysis," *Springer Computational Intelligence for Sentiment Analysis*, vol. 1049, pp. 312-322, 2023, doi: 10.1007/978-3-031-23028-8_8.

[12] A. Das and M. Barman, "Hidden Markov Models in Social Media Sentiment Analysis," *IEEE Transactions on Computational Social Systems*, vol. 8, no. 1, pp. 54-65, 2021, doi: 10.1109/TCSS.2021.3059196.

[13] T. Nguyen et al., "HMM-Based Sentiment Analysis for Product Reviews in E-Commerce Platforms," *ACM Transactions on Intelligent Systems and Technology*, vol. 13, no. 2, pp. 123-145, 2022, doi: 10.1145/3526178.

[14] Z. Li et al., "Improved Hidden Markov Model for Sentiment Analysis with Contextual Features," *Knowledge-Based Systems*, vol. 230, pp. 1-12, 2022, doi: 10.1016/j.knosys.2022.107562.

[15] B. Silva and A. Kumar, "An Analysis of Hidden Markov Models for Multi-Class Sentiment Classification," *Journal of Machine Learning Research*, vol. 21, pp. 1234-1256, 2023, doi: 10.1007/s10994-023-06123-8.