Hidden Markov Model for Cardholder Purchasing Pattern Prediction

Okoth Jeremiah Otieno, Michael Kimwele, Kennedy Ogada
Department of Computing, Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya

***Abstract*—This study utilizes the Hidden Markov Model to predict cardholder purchasing patterns by monitoring card transaction trends and profiling cardholders based on dominant transactional motivations across four merchant sectors, i.e., service centers, social joints, restaurants, and health facilities. The research addresses shortfalls with existing studies which often disregard credit, prepaid, and debit card transactions outside online transaction channels, primarily focusing only on credit card fraud detection. This research also addresses the challenges of existing prediction algorithms such as support vector machine, decision tree, and naïve Bayes classifiers. The research presents a three-phased Hidden Markov Model implementation starting with initialization, de-coding, and evaluation all executed through a Python script and further validated through a 2-fold cross-validation technique. The study uses an experimental design to systematically investigate cardholder transactional patterns, exposing training and validation data to varied initial and transition state probabilities to optimize prediction outcomes. The results are evaluated through three key metrics, i.e., accuracy, precision, and recall measures, achieving optimal performance of 100% for both accuracy and precision rates, with a 99% on recall rate, thereby outperforming existing predictive algorithms like support vector machine, decision tree, and Naïve Bayes classifiers. This study proves the Hidden Markov Model's effectiveness in dynamically modeling cardholder behaviors within merchant categories, offering a full understanding of the real motivations behind card transactions. The implication of this research encompasses enhancing merchant growth strategies by empowering card acquirers and issuers with a better approach to optimize their operations and marketing synergies based on a clear understanding of cardholder transactional patterns. Further, the research significantly contributes to consumer behavior analysis and predictive modeling within the card payments ecosystem.**

***Keywords*—Hidden Markov Model; cardholder transaction patterns; merchant categories; predictive algorithms**

I. Introduction

A. Introduction

This section details the scope of this research, an overview of the study in alignment with research objectives and the problem statement, brief definitions and literature on the use of the Hidden Markov Models, studies conducted at home and abroad, and finally, quick summaries on the importance of this research considering the recent works and gaps to be addressed.

B. Background

Cardholders show varied motivations prompting their transactional behavior, inclined to factors such as personal preferences, spending habits fashioned by income and financial status, demographics including age groups with separate spending patterns, and considerations of credit limits and debts. Transaction patterns also exhibit seasonal and temporal trends, with elevated spending levels during holidays and back-to-school periods, cognizant of prevailing economic conditions. These dynamics are essential to card payment service providers as they leverage them for optimizing cardholder spending and associated returns across different merchant categories and sectors.

The main objective of this research is to develop and evaluate a Hidden Markov model to predict the purchasing patterns of cardholders, aiming to address the failures of existing predictive algorithms such as support vector machines, Naïve Bayes classifier, and decision trees which have completely failed to offer an optimal prediction model due to their inherent challenges of lack of optimal generalizability. Further, these classical prediction algorithms have proved ineffective in several ways, for instance; the Naïve Bayes algorithm will usually not perform well when a 'train-test' approach is required since its assumption of feature independence is seldom true. Similarly, the Naïve Bayes algorithm is quite inefficient in handling unstructured datasets. On the other hand, decision trees are susceptible to overfitting, and they must be pruned to minimize the complexity parameter that controls the tree size and cut down on implying that slight variations in the training dataset may result in great differences in the trained hypothesis function, [12]. Lastly, support vector machines often exhibit long training time especially when dealing with large datasets and cannot stand alone in case of fault classification and when there is a necessity to distinguish between fault and no-fault while dealing with sparse data, [13]. In addition to addressing these challenges, this research presents a solution that tackles what most studies have completely overlooked. For instance, [32], [10], and [11] among others have revealed that predictive works only focus on online credit card user pattern prediction especially on fraud and associated risks, ignoring other cardholder behavioral patterns informed by intent, and other card types, and transaction channels such as debit, prepaid cards and card present usage channels that largely constitute volumes in card acceptance space.

C. Background and Literature Review

This research focuses on the Hidden Markov Model's predictive ability to profile cardholder purchasing patterns from merchant sectors and effectively assign a unique pattern based on dominant transactional trends. This research relies on the Hidden Markov Model' forward, backward, and expectation maximization algorithms to learn, update, and optimally predict the purchasing pattern of cardholders transacting at selected merchant categories representing the global card payment ecosystem. Besides online credit card transactions, this research scopes all card types across two main eligible card payment channels, i.e., card-present and e-commerce.

According to [1], the Hidden Markov Model is an augmentation of the Markov chain that describes probabilities associated with random variables, particularly observable events. The hidden Markov Model monitors both observed and hidden states. Further, [2] allude that Hidden Markov Model is useful in instances where data entails a sequence of observations, which are probabilistically independent of the internal state of a dynamical system.

Scholars have explored the predictive capabilities of the Hidden Markov Model in varied disciplines. These include predicting user behavior through data profiling, [2], driver behavior prediction based on environmental observation, [3], prediction of consumers' adoption behavior of products with water efficiency labeling, [4], enhancing credit card fraud detection, [5], risk assessment in cryptocurrency portfolios, [6], Motion Sequence Analysis, [7], spatial analysis of three-dimensional mineralization distribution, [8], and wavelet-based feature extraction [9] among other disciplines. This research explored the Hidden Markov Model's predictive ability by profiling cardholder purchasing patterns from merchant sectors.

In Kenya, scholars such as [31] have explored cardholder purchasing pattern prediction by focusing on anomaly detection, utilizing a hybrid approach that combines Hidden Markov Model with other machine learning algorithms. Further, [33] have used the Hidden Markov Model to explore the process mining techniques to detect fraud in banks, with particular focus on credit card behavior among other transactional patterns. From the two studies, it's evident that local scholars only focus on fraud detection on credit cards, while overlooking other patterns that inform transactional motivation, with a particular bias on product and channel being credit card and e-commerce respectively.

In the United States, scholars such as [26] explored cardholder purchasing patterns with a particular interest in online transaction fraud detection, utilizing the Hidden Markov Models. Similarly, in Southern Asia, scholars like [30] have explored the use of the Hidden Markov Model in predicting cardholder purchasing patterns, with a particular focus on credit card Fraud detection. A similar trend is reflected in West Africa, where [27] focused their research on credit and debit card fraud detection on Automated Teller machines through the Hidden Markov Model. Similarly, in Mexico, [11] utilized the Hidden Markov Model to explore cardholder purchasing patterns with a specific focus on fraud detection and identification in credit card transactions done via online channels.

In summary, the recent studies on the use of the Hidden Markov Model for cardholder purchasing pattern analysis suffer from an apparent bias in scope. Studies conducted within the Kenyan card payment and financial ecosystem, which are mirrored elsewhere, i.e., in the United States, West Africa, and Southern Asia fundamentally focus on fraud-based pattern prediction within the card-not-present payments ecosystem. Notably, these studies prioritize identifying fraudulent transactions only, overlooking the underlying cardholder intention behind genuine transactions, thus depicting an inadequate scope of intent. Additionally, the studies' exclusive focus on credit card transactions ignores other card types, i.e., debit and prepaid cards. These additional card types could potentially require tailored behavioral modeling that these studies completely ignore. Lastly, the studies exhibit utter channel bias, with the exploration of e-commerce-only transactions, ignoring the possibility of obtaining optimal pattern prediction results on other channels such as card-present point-of-sale terminals, till-integrated pin pads, and mobile point of sale that jointly present unique challenges and opportunities.

Against the backdrop of these notable inefficiencies, this research presents a novel approach of incorporating cardholder intent by monitoring dominant transactional patterns from merchant categories where cards are frequently run. This is done by profiling cardholders when they exhibit frequent appearances at specific merchant categories, associating their transactional preferences to the services offered at the said categories. Secondly, this study presents an expanded scope of all card types including debit and prepaid cards whose behavioral patterns are rarely explored by most recent studies. Lastly, this study embraces a multi-channel approach, scoping card-present and card-not present transaction channels to fully understand the effectiveness of Hidden Markov Model in cardholder purchasing pattern and transactional analysis. In summary, this study contributes significantly to the knowledge body by merging innovative statistical modeling capability of Hidden Markov Model with sector-specific transactional analysis to predict and comprehend cardholder purchasing patterns. Its procedural precision, novel approach to modeling cardholder purchasing pattern, and practical applications make it a remarkable addition to both academic world and the card payments industry.

II. Related Work

A. Introduction

This section offers an exhaustive review of related literature, emphasizing recent studies concerning the application of the Hidden Markov Model for predicting cardholder purchasing patterns. Additionally, it analyzes various existing prediction algorithms, including support vector machines, naïve Bayes classifiers, and decision trees, and specific works around these algorithms.

B. Existing Prediction Models

Besides Hidden Markov Model, this research explores the application of existing supervised machine learning algorithms for predicting cardholder purchasing patterns. These algorithms include support vector machines, Naive Bayes classifier, and decision trees. By assessing the performance of these algorithms, the research seeks to establish the most effective approach for predicting cardholder purchasing patterns.

According to study [16], support vector machines are a set of supervised and non-parametric statistical learning algorithms applied in regression, outlier detection, and classification. Support vector machines are used in problem classification in machine learning, aimed at generating the optimum decision edge that can separate n-dimensional space into classes for correct data point categorization in the future. The research in [17] used two steps to describe the working mechanism of the support vector machine. In their descriptions, the first step creates hyperplanes that separate the classes optimally. The second step chooses the right hyperplane with the optimum segregation from the nearest data points.

Recently, support vector machines have been used in cardholder purchasing pattern prediction. A typical instance is [18] who applied support vector machines in predicting e-commerce card customer-purchasing patterns and churn prevention. The scholars reviewed the main motivations for customers to make online purchases and reasons for attrition. The researchers acknowledged the possibility of predicting the customers' future inclinations using the necessary data acquired and the requisite analysis. The scholars used labeled datasets to train the support vector machines, giving them a reliable ability to predict and identify outcomes. However, the scholars noted that with the rise in e-commerce-based companies and clients, there was a need to automate the system to depict the results via a desktop application. To predict customer turnover in business-to-customer e-commerce, the scholars adopted a pronged model. The first approach employed SVM technology to foresee churn occurrences, while the second approach employed a hybrid strategy that blends collaborative, content-based, knowledge-based, and demo-graphic methodologies to derive customized retention approaches. The scholars concluded that computing the value of lost clients implied that as the count and rate of transactions grew, the risk of customer turnover considerably dropped. They further added that owners of e-commerce establishments should focus on the two factors of purchase volume and frequency, settle on a strategy for client retention, effectively limit customer attrition, and attain sustainable business development. Though the scholars achieved satisfactory outcomes with 77.36% in prediction accuracy from training the support vector machines, they acknowledged the long training time required by support vector machines especially when handling large datasets.

Similarly, [19] described a Bayesian classifier as a statistical technique of computing probability that a feature belongs to a class with the application of Bayes' theorem. The scholars added that Naive Bayes works with an assumption of in-dependence among predictors, can handle large datasets, and is equally easy to build with these two parts making up this algorithm: Naïve and Bayes. Researchers have recently explored the capabilities of Naïve Bayes, especially in card customer pattern predictions. For instance, [20] modeled binary classification of customers' online purchasing patterns using the Naïve Bayes classifier among other machine learning techniques. The paper proposed a method that utilized a subset of attributes based on the predictive ability and evaluated the duplication among selected traits, after which Naïve Bayes was applied. Using Naïve Bayes and other algorithms, the scholars used a hybrid approach that combined multiple classifiers to form an ensemble learner, to improve the classification outcomes. The experts used two datasets, namely Turkish bank, and German Credit datasets, through a feature selection stage. Though the study posted impressive results with predictive performance being 97.8% in precision, 88.2% in accuracy, and 89.1% in recall rate, the proposed approach still showed an overdependency on proper feature selection strategy, proving that Naïve Bayes assumption of feature independence makes it just unsuitable for real-world applications.

Correspondingly, the study in [22] defined a Decision tree as a machine-learning algorithm with nodes representing a predictor variable, the links between the nodes representing decision, and the leaf nodes representing outcomes or response variables. The scholars added that Decision trees are useful in splitting large datasets into smaller classes. Like any other machine-learning algorithm, decision trees have been used in card customer purchasing pattern prediction. One notable instance is that of [29] who used decision trees to predict cardholder purchasing patterns with a focus on fraud detection and prevention. The study proposed a theoretical credit card fraud detection and prevention model that used a phased structure from initiation to implementation. The scholar began with data collection, model filtering, and model selection for the implementation through the decision tree framework. The scholar used datasets from Kaggle that had several credit card transactions, with a complimentary binary field indicating the flag, whether a transaction is fraudulent or genuine. The datasets used equally included additional features such as trans-action amount, transaction time, transaction frequency, merchant information, geographical location, cardholder information, transaction type, and device type and in-formation. Secondly, the study filtered data and applied a decision tree classifier algorithm to perform credit fraud detection and prevention utilizing common features. The decision tree classifier algorithm started with the root node and then divided the data components into internal nodes to depict dataset features. The decision tree classifier algorithm mapped the features into decisions and the results. The decision tree classifier model posed a query and then divided this into sub-trees in line with corresponding responses. The model used by the scholar started its search at the top of the tree to locate the dataset classes. The model then circulated the branches based on the related and matched the base trait with the record attribute to arrive at the next node. The de-signed model worked on the user's device as a client-side implementation module to forecast the possibility of fraud in credit card transactions. Further, the generated forecasts would trigger real-time feedback to the user regarding the transaction risk level. The author evaluated the model through sensitivity and precision which delivered a performance level of 81.6%.

C. Hidden Markov Models

Recently, scholars have explored the capabilities of Hidden Markov Models in several predictive works. For instance, [25] explored the capabilities of Hidden Markov Models in predicting the behavior of e-commerce card customers to project store income. The study factored three items in the final predictive model, i.e., loyalty, vendor, and psychology which returned the probabilities utilized in the transition matrix of the hidden Markov model, giving three decision-states, i.e., 'Order completed', 'Order uncompleted, or 'No order'. The model outputs were interpreted by the Viterbi algorithm to approximate if the order has been completed successfully, followed by the evaluation of the predicted store income. The authors base-lined their model to the prediction presented by the Google Analytics tracking system. While considering the sub-models, the researchers simplified the vendor sub-model by taking into account only two vendors' market share without 100% domination, if 100% monopole is quite impossible in a real market. The Vendor sub-model returned the probability of a customer sticking to one vendor J, and not shifting to another vendor K. The researchers denoted this by [0,1], with 0 implying that a customer changed to vendor K, while leaving store J without fulfilling an order, whereas 1 means a customer stuck to store J, thus fulfilling the order. In considering psychology, the scholars alluded to the role of psychology as a behavior stimulant in situations when the customer gets influenced. The scholars mentioned societal pressure, price differences, mood aspect, the center of mass effect, and actual need among others as key factors that drive the psychological sub-model. The study picked the price aspect and center of mass effect, both representing the probability of a customer's resolve to make a purchase. The study viewed the loyalty aspect as the customers' positive feeling towards a store, and the decision to stick to that store no matter what happened. The study segmented loyalty into two components, namely, fidelity and commitment. The study evaluated the performance of the two models using the R2 technique and the criteria revealed the Hidden Markov Model's superiority as it outperformed the Google Analytics tracking system with Hidden Markov Model posting 0.95 R2 performance compared to Google Analytics Tracking System's 0.90 R2 Performance.

In another instance, [26] highlighted the performance of the Hidden Markov Model while predicting the behavior of credit card transactions on e-commerce channels. The study indicated that online card transactions are more convenient as customers do not have to visit the stores. However, the study also indicated that online card transactions are a prime target for fraud due to the nature of virtual requirements, i.e., card number, card verification value, and expiration date. The study used the Hidden Markov Model to detect fraud at the time of the transaction, applying blocking mechanisms that could bar any flagged transactions. The study also used behavior analysis to understand the spending patterns of the cardholders. The study revealed the capability of the Hidden Markov Model to acquire a high-level fraud analysis with a minimal false alarm ratio. The study's model was able to predict the behavior through three price ranges, i.e., low, medium, and high, depending on the loaded transaction amount. Further, the model-built clusters depend on the price ranges. Depending on the cardholder spending trend, the model figured out any variations noticed. The study settled on initial state probabilities based on the previous data which were structured to make a sequence of future forecasts. Through the stochastic process of the Hidden Markov model, the study built a predictive system that attained 80% accuracy and precision levels and established the Hidden Marko Model's suitability to model a fraud detection system in an online transaction system. Furthermore, the model constructed proved efficient in handling large amounts of transactions and giving results swiftly.

Further, [27] reviewed the detection of fraud in automated teller machine transactions using the Hidden Markov Model. The study used three model parameter estimation approaches considering the model and the observation sequence, i.e., pattern recognition problem which computed the probability of an observed sequence of a given Hidden Markov Model, labeled evaluation problem, and computing the sequence states responsible for the observation sequence, christened decoding problem, and finally, generating a sequence of observations, which was christened as learning problem. The scholar constructed an automated teller machine fraud detection model that baselined on cardholders' spending habits. This approach sampled three different cardholder spending profiles depending on the price range, which was categorized as Low, Medium, and High. The lowest price was set at zero, with a specific amount as medium price, and finally the card limit as the highest price range. The prediction was systematically done in two phases, i.e., the training phase which involved the initialization of Hidden Markov model parameters, followed by iterative estimations of forward and backward steps, and finally, the detection phase which involved the Hidden Markov Model verifying the fraudulent transactions using a clustering algorithm. The model developed had considerably better performance as measured through a sensitivity value of 85.5%, in comparison to the Gaussian mixture model which gave a sensitivity of 40%.

In another instance, [15] examined detecting electronic banking fraud on highly imbalanced data using hidden Markov models. The authors proposed a framework based on the Hidden Markov Model, modified density-based spatial clustering of applications with noise alongside synthetic minority over-sampling approach to detect fraud in a highly imbalanced electronic banking dataset, the bulk of which involved credit card transaction data. The authors considered some transactional attributes while building the model, i.e., transaction amount, transaction frequency, and transaction types. The scholars utilized datasets from the Kaggle public repository to conduct the research. The dataset had a variety of parameters around the transaction attributes, for example, transfer, cashout, cash-in, debit, and payments, with a general categorization of genuine transactions and fraudulent transactions. The scholars categorized the transaction amounts as either low, initialized by l= (0,100), medium, initialized by m= (100,500), and high, initialized by h= (500, Transaction Limit) values. The study also regarded the frequency of transaction occurrence as either low, indicated by transactions happening less than five times a month, intermediate, depicted by transactions happening between 5 and 10 times a month, and high, signaled by transactions happening at least 10 times a month. The authors regarded various transaction types as internal states while treating transaction amounts and frequency as observation symbols. The study performed four sets of simulations in two stages using Python and compared the performance. The study produced fairly good results, with performance recorded as follows, precision rate was 95%, recall rate was 0.97, and accuracy of 95%.

In a separate study, [30] explored the application of the hidden Markov model in the study of fraud detection systems where they acknowledged the most recent events such as the Coronavirus pandemic that compelled the world to systematic adoption of contactless payments, especially credit card payments. The research equally acknowledged that with the increasing adoption of credit card payments, fraudulent transactions are increasing, with fraudsters inventing new techniques to keep defrauding issuers and acquirers. The scholars modeled a system that checks credit card transactions online, profiling cardholders based on whether they are authorized or not. When an authorized cardholder does a transaction, their profile is matched on the backend database and if the match is successful, the transaction is processed success-fully, notifying the user of success. Equally, if an unauthorized card performed a transaction and their spending profile failed to match what exists in the back.

Recently, scholars have explored the use of Hidden Markov Models in customer purchasing pattern predictions but with numerous gaps. For instance, [21] affirmed that most scholars get unreliable results because they use relatively minimal datasets that deprive the research of comprehensibility of the results for testing and resilience. Moreover, the study alluded that minimal preprocessing could lead to un-balanced data which increases the likelihood of bias in the model creation. Norah also confirmed that many researchers exploring the capabilities of the Hidden Markov Model fail to attain reliable results because of the limited number of iterations which makes it impossible to attain optimal outcomes. A further review of recent works on hidden Markov models revealed notable gaps especially when it comes to the type of datasets used. For instance, a study by [23] revealed that utilizing extremely historical data may not give the correct prediction especially where parameters such as perception and customer behavior change over time. In addition to this, customer pattern prediction hasn't been reliably done because most scholars overlook the demographics and assume that the personalities and attributes of customers are universal, ignoring age, and geographical location among others. The research in [23] also confirmed that most research works suffer from overfitting, where training data gives excellent results while real data speaks otherwise. Correspondingly, the study in [24] confirmed that most research works get unreliable results because of the process of data collection which doesn't give assurance of trust and accuracy. The scholars acknowledged the need to improve the data collection process to ascertain more precise and comprehensive data that plays a major role in improving model efficiency. In addition, the study in [28] explored the application of the hidden Markov model for credit card fraud detection. However, their work had notable gaps, especially in the lack of clarity of the performance metrics. With claims of the model achieving high accuracy and optimal processing speed, the scholars did not attempt to quantify the performance metrics in absolute figures, making it difficult to tell the actual performance of the model. Like other scholars reviewed in this research, the model only focused on credit card fraud detection for online transactions, an area of saturation in cardholder purchasing pattern prediction, ignoring other behavioral aspects related to cardholder purchasing patterns.

III. Methodology

A. Introduction

This section highlights the data acquisition and preprocessing, design methodology, materials and methods used, and ethical considerations for this research.

B. Data Acquisition and Pre-processing

This Study explores the application of the Hidden Markov Model to predict cardholder purchasing patterns, with an interest in sector-based pattern analysis. The research applied random sampling to select card transactions relating to specific sectors. Random sampling was useful as it allowed unbiased data collection, qualifying our re-search to arrive at fair and unbiased conclusions. Additionally, this study considered random sampling due to its efficient generalizability of research findings, and proper statistical inferencing. The data was extracted through structured queries written in HiveQL, using the Hadoop Ambari workbench. Only selected columns were queried from the transaction tables then upon query execution, the data was exported to a .csv file for ease of manipulation. The data contained 4,500 records with the following fields: Card Number, which remained masked due to data privacy policy, Source, Merchant Name, Transaction date, Merchant Category Code (MCC), and Sector. The sector field was the most crucial as it dictated the initial state probabilities for each cardholder. The four sectors included restaurants, social joints, health facilities, and service centers. The sample dataset was captured in Table I below.

TABLE I. SAMPLE WORKING DATA

Card Number	Merchant Name	Date	Sector
****6840	Merchant 1	Apr-21	Service Centers
****0010	Merchant 2	Apr-21	Restaurants
****2540	Merchant 3	Apr-21	Social Joints
****9110	Merchant 4	Apr-21	Service Centers
****1080	Merchant 5	Apr-21	Restaurants
****0200	Merchant 6	Apr-21	Social Joints

The data used in this research was cleaned to remove noises, unnecessary columns, and outliers through R studio and Python's Jupyter Notebook, all packages within the Anaconda ecosystem. For ease of manipulation, this research applied feature engineering to create binary variables that expressly indicate specific merchant categories and related sectors. Based on the specific merchant categories, a set of states was defined to represent the intent of cardholders as they visit the merchant stores to do transactions.

C. Design Methodology, Methods and Tools

This research used experimental design methodology to ratify the suitability of the Hidden Markov model in predicting cardholder purchasing patterns. The approach employed the Hidden Markov Model's computational modeling for optimal prediction. The study is baselined on the following connotations and assumptions of the Hidden Markov Model highlighted by [1], majorly on the shorthand equation:

$$\lambda = (A, B, \pi) \tag{1}$$

1) General notations

- N as the number of states in the Hidden Markov Model.
- M as the number of observation symbols.
- T as the length of the observation sequence.

- π as the initial state probabilities (N×1 vector).
- A as the transition matrix (N×N matrix).
- B as the emission matrix (N×M matrix).
- α as the forward variable (T×N matrix).
- β as the backward variable (T×N matrix).
- γ as the state occupation probability (T×N matrix).

2) Forward algorithm: The forward algorithm calculates the probability of observing a sequence of symbols given the Hidden Markov Model.

$$\alpha_t(j) = \sum_{i=1}^{N}\left([\alpha_{t-1}(i)\, A_{ij}] \times B_j O_t\right) \quad (2)$$

3) Backward algorithm: The backward algorithm calculates the probability of observing the rest of the sequence given the current state.

$$\beta_t(i) = \sum_{j=1}^{N}\left(A_{ij} \times B_j \text{ x } B_j(O_{t+1}) \times B_{t+1}(j)\right) \quad (3)$$

4) Expectation step: The Expectation step computes the probability of being in a particular state at a particular time given the observed sequence.

$$\gamma_t(i) = \frac{\alpha_t(i) \times \beta_t(i)}{\sum_{j=1}^{N} \alpha_t(j) \times \beta_t(j)} \quad (4)$$

5) Maximization step: This step updates the model parameters (π, A, B) using the γ values obtained in the e-step.

Update Initial Probabilities: $\pi_i^{New}\gamma_i(i)$ (5)

Update Transition Matrix:

$$\mathrm{A}_{ij}^{New} = \frac{\sum_{t=1}^{T-1} \gamma_t(i) \times A_{ij} \times B_j(O_{t+1}) \times B_{t+1}(j)}{\sum_{t=1}^{T-1} \times \sum_{t=1}^{T-1} \gamma_t(i) \times A_{ij} \times B_j(O_{t+1}) \times B_{t+1}(j)} \quad (6)$$

Update Emission Matrix:

$$B_j(k)^{New} = \frac{\sum_{t=1}^{T} \gamma_t(j) \; if \, O_t = k}{\sum_{t=1}^{T} \gamma_t(j)} \quad (7)$$

6) Convergenece: The EM algorithm iteratively performs E-step and M-step until convergence. Convergence is determined by assessing if the change in model parameters is below a certain threshold. The study in [14] described the three fundamental problems that Hidden Markov Model seeks to address: a) Evaluation Problem: This always requires summing over all possible hidden state variables. b) Decoding Problem: Prompts us to compute the best sequence of hidden states, and c) Learning Problem or Optimization Problem: Evaluate an observation sequence (O1, O2…On and the Hidden Markov Model λ = (A, B, π) that optimizes the probability of O.

Further, this study followed the below approaches to fulfill experimental design methodology and tailor the desired output:

a) Data partitioning: The dataset was partitioned into two folds, with fold 1 acting as training data while fold 2 as validation data. The training dataset was used to construct the Hidden Markov Model, while the validation dataset was used to perform hyperparameter tuning. The two partitions were as per Table II below:

TABLE II. DATA PARTITIONS

Partition	No of Records
Fold 1	2,500
Fold 2	2,000

b) Benchmark model: The study applied two separate attempts to train the model and the first attempt's outcome was used as the baseline model against which the subsequent results were evaluated, c) Performance Comparison: The performance of the Hidden Markov Model for future prediction of cardholder purchasing was compared against the baseline model. Further, the performance of the resultant model was compared with models built on support vector machines, decision trees, and Naïve Bayes classifier. Performance Metrics like precision, accuracy, and recall were essential for this comparison, and lastly, d) Sensitivity Analysis: This research varied the initial state probabilities to assess the performance of the resultant model under different circumstances and datasets.

In summary, this research used the below simple steps to fulfill the implementation process:

- Initialization: Defined the initial state probabilities (π) based on the cleaned dataset distribution.
- Transition and Emission Probabilities: Matrices (A and B) representing probabilities of state transitions and emissions were randomly initialized.
- Observation Sequence: Random sequence of observations was generated to simulate card transactions across sectors.
- Model Training: This study implemented Expectation Maximization algorithm to adjust the parameters (π, A, B) iteratively over 100 iterations until convergence was attained.
- Model Validation: This study used a 2-fold cross-validation approach to validate the model performance, comparing against other algorithms such as Support Vector Machines, Naïve Bayes, and Decision Trees for accuracy, precision, and recall rate.

D. Ethical Considerations

This research scoped the following ethical considerations during its formal execution:

1) Data privacy and confidentiality: - Since this research focused on sensitive credit, debit and prepaid card transactions data, all sensitive details were handled with utmost privacy and confidentiality. Card numbers were first tokenized then subsequent tokens masked to curtail any compromise. Furthermore, this research did not expose the card verification values and the expiration dates in congruence with the non-disclosure agreement signed with the data provider.

2) Informed consent: - Customers whose details were used in this research received consent notification via short message service, given options to opt out if not comfortable with the process, assuring them of the security and confidentiality of their data. Customers that replied with an 'opt-out' had their details expunged from the dataset.

3) Fairness and equity: - This research treated all sampled participants with utmost fairness regardless of their demographic characteristics. All sampled participants had an equal opportunity to contribute to the findings of this study.

4) Regulatory compliance: - This research adhered to the necessary regulations governing collection, storage, and utilization of card data in alignment with the local regulator and payment card industry data security standards.

IV. Hidden Amrkov Model for Cardholder Purchasing Pattern Prediction

A. Introduction

This section highlights the step-by-step model construction, required parameters, full implementation, and architecture in context.

B. Background

Credit, debit and prepaid cardholders usually visit merchant stores through different channels to fulfil their purchasing needs by paying via card in exchange. Different channels enable merchants to accept card payments, i.e., card present channels where customers must physically visit the merchant stores to run their cards on a point-of-sale terminal issued by an acquiring bank. Additionally, merchants can also do transactions through card not present channel (electronic commerce) where they do not have to be physically present at a merchant store but can shop over the internet and make arrangements to receive the goods or services upon payment. Several factors are key determinants and motivations to which store a cardholder would visit, and these factors include the following: income level, personal preference, location and proximity, store loyalty programs, product selection and quality, seasonality, market trends and economic conditions, online shopping habits among others.

This study focused on cardholder purchasing pattern from a standpoint of sectors frequently visited for transaction as these would inform the behavioral patterns mentioned earlier. In simple, this study treated key transactional motivations as the main pointers to a definite cardholder purchasing pattern. In a normal card payments eco-system, cardholders can have various motivations for doing transactions, with each motive informed by an underlying need, which could be either be genuine or malicious, intent to do fraud and other genuine reasons.

Further, the study reviewed all card types and categorized the purchasing patterns by the following sectors: Service Centers (S), Social Joints (J), Health(H) and Restau-rants (R). Each sector was marked with an observation sequence treated as customer motivation, being the desired purchasing patterns. The intentions were mapped to each sector as per Table III below:

TABLE III. Observation State Symbols

Initial State Symbols	Sector Represented	Observable States	Symbol
H	Health	Auto service	A
J	Social Joints	Medication	M
R	Restaurants	Eating	D
S	Service Centers	Entertainment	H

At any point in time, a cardholder running their card at a given sector outlet automatically qualifies for an observation state associated with that sector. A card that is run on a platform manned by a service center(S) is deemed to adopt a pattern that leans towards auto service and is assigned state (A), a card that is run at a restaurant automatically exhibits a pattern relating to eating (E) and is assigned state (E), while cards that are run at social joints (J) and health care (H) are associated with patterns entertainment (H) and medication (M) respectively. This mapping is shown in Fig. 1 below:

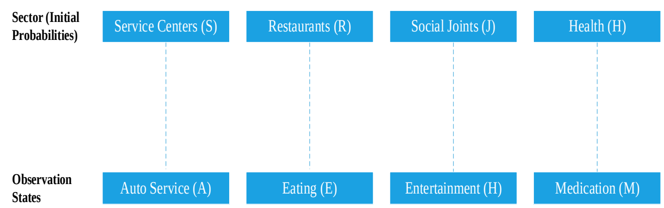

Fig. 1. Observation state mappings.

In the context of our study, observation states A, E, H and M are always hidden, and these are the pointers to unravelling the intention or purchasing patterns of cardholders of cardholders as they go about doing transactions at different outlets.

C. Purchasing Pattern Prediction

In this study, we implemented the cardholder purchasing pattern prediction using a python script that worked through three phases, i.e., initialization phase, decoding phase, and evaluation phase.

The first phase of initialization involved defining the initial state probabilities in alignment with Fig. 1 above to be in a particular state, with states being retained as S, J, R and H. The initial state probabilities were computed from the transactional distribution on the dataset. Considering the distribution of each state on the dataset, the initial state probabilities were mapped as follows: S, J, R, H with each corresponding to the items in the matrix π = [0.1148, 0.292, 0.4, 0.5532]. Secondly, the transition probability between states were randomized by the following 4 by 4 matrices to represent the transition between one state to the next, denoted by A:

$$A = \begin{matrix} 0.1 & 0.3 & 0.5 & 0.1 \\ 0.2 & 0.3 & 0.1 & 0.4 \\ 0.6 & 0.1 & 0.1 & 0.2 \\ 0.1 & .03 & 0.3 & 0.3 \end{matrix}$$

Thirdly, emission probabilities were also randomized using the below 4 by 4 matrices to depict the possibilities of emissions from each state, denoted by B as below:

$$B = \begin{matrix} 0.2 & 0.3 & 0.4 & 0.1 \\ 0.4 & 0.2 & 0.1 & 0.3 \\ 0.3 & 0.3 & 0.2 & 0.2 \\ 0.1 & .02 & 0.3 & 0.4 \end{matrix}$$

Next, the observation sequence was randomly generated as ([0, 1, 2, 3, 2, 1]) to represent the sequence of observed emissions. Lastly, the expectation maximization function algorithm was defined, initializing parameters randomly and performing iterative updates to obtain optimal values. The maximum iteration count was set at 100.

The second phase, i.e., decoding involved parameter initialization of the Hidden Markov Model problem function, λ = (A, B, π), where λ corresponded to the probability of being in each state at each time step given all the observations, A corresponded to transition matrix, B corresponded to emission matrix, and finally π corresponded to initial state probabilities. This was followed by a forward algorithm which involved systematic computation of the probability of being in each state at each time step provided the observation up to that step. Further, a backward algorithm was performed to compute the probability of observing future emissions given the present state. Both forward pass and backward pass were sequentially performed by a for loop iteration in the python script, with forward pass denoted by alpha (α) and backward pass denoted by beta (β). After optimizing alpha (α) and beta (β), the probability of being in each state at each time step given all the observations was computed, denoted by gamma (γ). After computing the λ, the joint probability of being in two states at consecutive time steps, denoted by (xi) was computed to close out the second phase.

The third phase, i.e., evaluation involved updating the parameters, convergence, and the final output. On updating the parameters, the model utilized the calculated gamma and (xi) to update the parameters of A, B and π of the Hidden Markov Model, through adjusting the initial state probabilities, transition probabilities and emission probabilities to better fit the observed sequence. At convergence, the model iterated 100 times, refining the parameters at each iteration until convergence. On running the expectation maximization algorithm, optimal transition and emission matrices were obtained.

In summary, this study followed the below procedure to adequately execute and update the model hyperparameters in line with the intended scientific contributions to the knowledge body:

- Initialization Update: we initialized the parameters (π, A, B) based on data distribution and postulates. In this research, π represents the initial state probabilities, with A and B initialized to reflect possible state transitions and emission probabilities.
- Iteration and Convergence: In this research, the expectation-maximization algorithm updates π, A, and B iteratively to optimize the likelihood of the observed sequence. This iterative process progresses until the convergence of parameters.
- Parameter Adjustment: At every iteration, π, A, and B are modified based on γ, and ξ calculated in the expectation step and maximization step. This modification polishes the model's estimates to fit the observed sequence of transactions better.

Essentially, the model trained Hidden Markov model on a given sequence of observations using expectation maximization algorithm by adjusting the model parameters to optimize the likelihood of observed data.

This study leverages the capabilities of the Hidden Markov Model to predict the purchasing pattern of cardholders with the following pillars that make it stand out within the scientific community:

- Excellent behavioral analysis and modeling: This research offers precise modeling and analysis of cardholder transaction patterns through a sector and motive-driven approach. This is achieved by iterative refinement of π, A, and B to predict future patterns based on past transactional behavior.
- Sector-Specific Hidden Markov Model: Unlike traditional hidden Markov models used for sequenced prediction modeling, this research presents a unique application of Hidden Markov Model to predict cardholder purchasing patterns by profiling cardholders based on sectors such as health, social joints, restaurants, and service centers, representing primary transaction motivations.
- Observation State Mapping: This study presents a distinctive mapping between the observed transactions at different sector outlets, for instance, running a card at a restaurant, and hidden states depicting the motivation behind the purchase such as eating. This level of mapping explains the reasons behind card usage.
- Applicability in the card payments industry: This research provides great insights into card transactional data, highlighting the dynamics of consumer spending guided by intent. This contributes to comprehending card payment dynamics. With this knowledge, card acquirers and issuers can perform personalized marketing on customers whose purchasing patterns have been identified over time. The model was implemented in three phases as per the sequence diagrams below in Fig. 2.

Further, the model was mapped to show full process flows, implemented in a full architecture diagram as depicted by Fig. 3.

In summary, this study focused on a hidden Markov Model based on a 2-fold cross validation with the first fold subjected to three sets of training, shifting, and changing the initial state probabilities to alter the status and monitor the performance under different circumstances. Further, the study involved the validation of the model using a second fold which was subjected to similar circumstance to perfect the model performance. The model borrowed from HMM's expectation maximization algorithm due to its inherent ability to offer interpretability owing to its explicit modelling of hidden states.

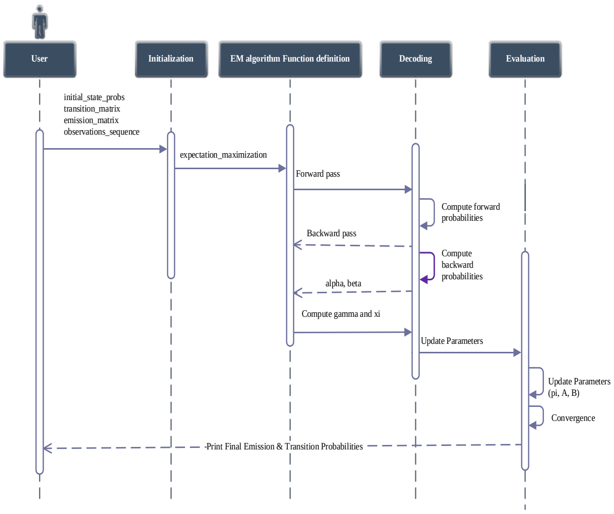

Fig. 2. Three-phased implementation sequence diagrams.

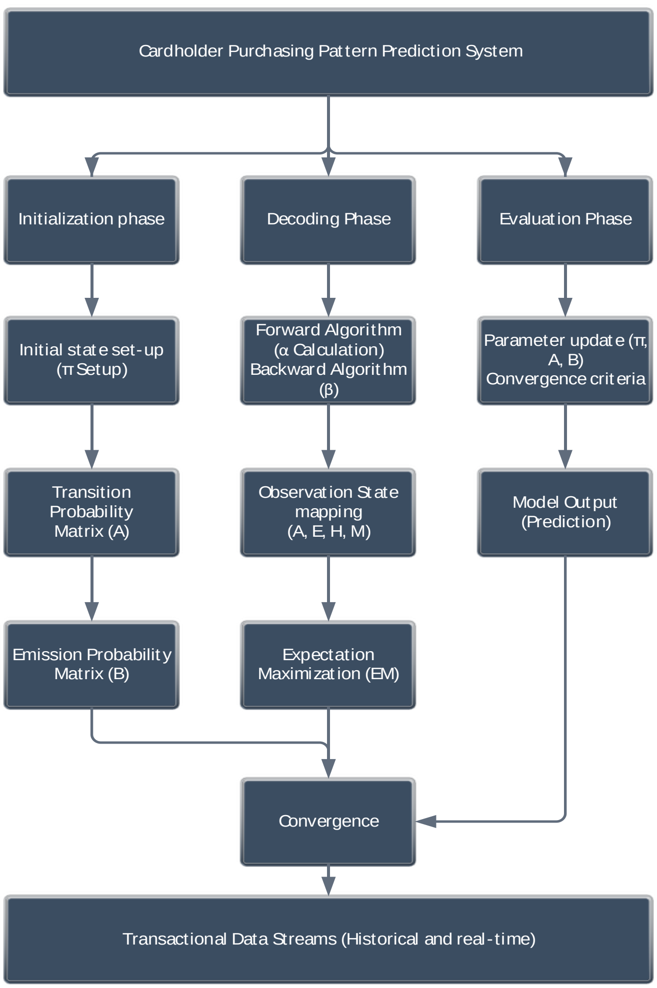

Fig. 3. Hidden Markov Model for cardholder purchasing pattern prediction architectural diagram.

V. RESULTS

A. Introduction

This section highlights the training and validation results, encompassing the Hidden Markov Model results and those of existing algorithms. Further, the section capture results comparison at different levels, and finally, the discussion after each set of results.

B. Model Training and Validation

The model was trained and validated using the first and second fold of our datasets respectively and the results were recorded as per Table IV and Table V below:

TABLE IV. BASELINE MODEL AND TRAINING RESULTS

First Attempt			
Initial State Probabilities			
0.1148	0.292	0.04	0.5532
Observation Array			
287	730	100	1383
Optimized Transition Probabilities			
1.31E-81	1.64E-80	6.93E-81	2.04E-76
2.24E-58	2.39E-57	1.22E-58	1.80E-52
5.38E-73	3.81E-73	9.55E-74	3.16E-68
7.16E-09	2.46E-07	2.39E-08	3.00E+00
New Observation sequence			
1383	1383	1383	625
Optimized Observation probabilities			
0.29	0.29	0.29	0.131
New Normalized Observation Array			
724	724	724	327
Second Attempt			
Initial State Probabilities			
0.292	0.1148	0.5532	0.04
Observation Array			
287	730	100	1383
Optimized Transition Probabilities			
2.73E-56	1.18E-59	8.38E-50	2.97E-60
4.08E-143	2.55E-145	7.43E-138	2.92E-145
3.85E-07	7.75E-12	3.00E+00	1.31E-11
5.49E-140	1.10E-142	9.92E-134	8.91E-143
New Observation sequence			
1383	1383	1383	625
Optimized Observation probabilities			
0.29	0.29	0.29	0.131
New Normalized Observation Array			
724	724	724	327

TABLE V. MODEL VALIDATION RESULTS

First Attempt			
Initial State Probabilities			
0.2215	0.108	0.0705	0.6
Observation Array			
443	216	141	1200
Optimized Transition Probabilities			
4.13E-70	1.95E-69	2.24E-69	3.08E-65
1.09E-58	2.31E-58	5.67E-59	1.58E-53
8.62E-62	1.96E-62	1.47E-62	1.94E-57
3.32E-08	2.08E-07	1.15E-07	3.00E+00
New Observation sequence			
1200	1200	1200	500
Optimized Observation probabilities			
0.293	0.293	0.293	0.122
New Normalized Observation Array			
585	585	585	244
Second Attempt			
Initial State Probabilities			
0.1	0.5	0.1	0.3
Observation Array			
443	216	141	1200
Optimized Transition Probabilities			
1.25E-85	1.42E-79	6.15E-85	1.00E-84
4.79E-09	3.00E+00	2.18E-09	2.30E-07
1.64E-89	5.58E-85	2.70E-90	3.03E-89
2.33E-59	5.15E-53	6.70E-59	8.28E-58
New Observation sequence			
1200	1200	1200	500
Optimized Observation probabilities			
0.293	0.293	0.293	0.122
New Normalized Observation Array			
585	585	585	244

The model was trained using the first fold in Table II, using randomized initial state probabilities while utilizing the cardholder distribution from the dataset as an initial observation array. The dataset was subjected to two training attempts with varied and randomly assigned initial state probabilities. The model estimated the optimal transition matrices per attempt and recorded the results alongside optimal observation sequences. The observation sequences were normalized through their new probability distribution, allowing the model to compute and estimate the new observation sequences. It was evident that with different initial state probabilities, the model proved adaptable, giving comparable results despite the frequent alteration of the initial state distributions through various iterations.

To confirm the consistency of our model, we changed the dataset and observed the patterns and the new predictions. The latest dataset was run through two attempts initial state probabilities. The results recorded had striking similarities with varied to the previous results obtained during the model training stage, with noticeable similarities across different parameters. For each set of optimized transition probabilities, we assessed model performance using a confusion matrix to obtain accuracy, computed the precision, and recall metrics. Both training and validation datasets exhibited outstanding performance, achieving 100% accuracy and precision, with a recall rate of 99%.

Our model proved efficient when subjected to training and test data under different circumstances, with varied initial state probabilities. In all the two attempts for both training and validation, the results posted an accuracy of 100% when evaluated through a confusing matrix, a recall rate of 99%, and a precision of 100%. Assertively, the distribution across the results remained diverse, with a converging efficiency.

Further, we ran our datasets through models built from support vector machines, Naïve Bayes and Decision trees with results recorded in Table VI.

TABLE VI. NAÏVE BAYES, SUPPORT VECTOR MACHINES AND DECISION TREES

Naïve Bayes
Optimal Results:
Predicted, Health, Actual, Social Joints
Decision Trees
Number of transactions: 600, Predicted Sector: S
Number of transactions: 300, Predicted Sector: J
Number of transactions: 1500, Predicted Sector: R
Support Vector Machines
Optimal transaction volumes of 600, Predicted sector: S
Optimal transaction volumes of 1200, Predicted sector: R
Optimal transaction volumes of 500, Predicted sector: S
Optimal transaction volumes of 700, Predicted sector: S

The results from Table VI were evaluated through the metrics and efficiency measures recorded in Table VII below:

TABLE VII. PERFORMANCE ANALYSIS OF NAÏVE BAYES, SUPPORT VECTOR MACHINES AND DECISION

Algorithm	Accuracy	Precision	Recall
Support Vector Machines	80%	80%	80%
Naïve Bayes	75%	75%	75%
Decision Tree	80%	80%	80%

To predict optimal patterns with Support Vector Machines, we noticed that support vector machines encountered limitations in decoding sequential data when predicting cardholder purchasing patterns due to their intrinsic characteristics and design. Though support vector machines recorded a consistently fair performance across all three metrics, the model struggled with encoding progressive dependencies and capturing the sequential nature of the data, giving lower performance ratings in comparison to the Hidden Markov Model. Further, we noticed that decision trees equally posted fair results across the three metrics. However, the performance was lower than the Hidden Markov Model due to Decision trees' tendency to make decisions based on static features at each node, deficiently capturing the dynamic attributes intrinsic in sequential data. Finally, Naive Bayes classifiers equally posted commendable performance across the three metrics, though with its fair share of shortfalls and sub-optimal efficiency compared to the Hidden Markov Model. This was due to the Naïve Bayes classifier's assumption that all features are independent given the class label that did not hold for our dataset, where the order and interdependencies between observations were crucial. In the context of cardholder purchasing patterns, the dataset contained temporal dependencies and correlations between transactions, which Naive Bayes failed to capture effectively.

C. Comparison with Related Works

We reviewed the performance of our model against similar studies conducted recently and the outcome recorded in Table VIII below:

TABLE VIII. COMPARISON WITH EARLIER STUDIES AND EXISTING ALGORITHMS

Author	Algorithm	Accuracy	Precision	Recall
Okoth et al. (2024)	Hidden Markov Model	100%	100%	99%
[25]	Hidden Markov Model	95%	95%	95%
[26]	Hidden Markov Model	80%	80%	80%
[27]	Hidden Markov Model	85.5%	85.5%	85.5%
[30]	Hidden Markov Model	85%	85%	82%
[10]	Hidden Markov Model	95%	95%	97%
[18]	Support Vector Machine	77.36%	77.36%	77.4%
[20]	Naïve Bayes Classifier	88.2%	87.8%	89.1%
[29]	Decision Tree Classifier	81.6%	81.6%	82%

From Table VIII, it was evident that our study shared similarities with other works in the following aspects: Across all the studies, hidden Markov models posted relatively higher accuracy levels ranging from 80% to 100% with [26] recording the least accuracy of 80% while our model recorded the optimal accuracy of 100%. Equally, the studies recorded consistent precision levels of between 80%-100%, proving the suitability of hidden Markov models in such datasets. Likewise, the recall rates were relatively high, with a consistent record of between 80%-99%, proving that the models created in different studies effectively captured the relevant data points across related datasets. Overall, the above similarities strongly suggested that the Hidden Markov Model implemented by various authors was generally robust and performed well across related datasets, attaining commendable accuracy, precision, and recall levels. Similarly, the performance of our model was also compared with the performance of the existing algorithms, against our dataset and Hidden Markov Model posted optimal results of 100% in accuracy and precision, and an outstanding recall rate of 99% while support vector machines and Decision trees posted 80% in accuracy and precision with a recall rate of 80% respectively. On the other hand, Naïve Bayes posted a performance of 75% in both accuracy and precision, with a 75% recall rate on our dataset.

In summary, our model proved to be unique in three aspects, i.e., it achieved perfect performance metrics, with accuracy, precision, and recall rates of 100%, 100%, and 99% respectively, indicating flawless performance in prediction and data classification. Whilst the other authors also attained commendable performance, their metrics varied slightly, with accuracies ranging from 80% to 95.4%, precisions from 80% to 95.4%, and recalls from 80% to 97%. This was proof that our model outperformed others in terms of these metrics. Additionally, our model stood out for its comprehensive consideration of all card types, i.e., credit, debit, and prepaid cards run on all available card acceptance platforms, i.e., card present point of sale terminals and card not present e-commerce, considering that other authors only scoped credit card transactions on e-commerce, restricting their works to the only card, not present credit card transactions. Additionally, most of the works considered in this study revealed that most researchers reviewed cardholder purchasing patterns with an assumption that card transactions can only happen online via credit card, or on automated teller machines via debit cards. Conversely, our study revealed that in a card ecosystem, there are different card acceptance platforms across card present (physical point of sale terminals and automated teller machines) and card not present (Online) channels that can accept payments from credit, debit, and prepaid cards. Thirdly, whereas most authors only considered credit card fraud as a dominant pattern for cardholder purchase intentions, our study scoped cardholder behavior at the transaction level, considering the dynamic shift between one sector to the other, covering the day-to-day behavioral patterns that might inform a card transaction. Using the patterns identified from the inter-sectoral transactions, card-holders were profiled based on transactional frequency around a given sector and critical decisions made at the business, and/or sector level. Considering these facts, our model proved versatile and inclusive as it scoped a wide range of cards and platforms and proved beneficial in multi-behavioral pattern prediction.

D. Research Significance and Key Pointers

This research is quite significant because it achieved excellent performance in predicting cardholder purchasing patterns using the Hidden Markov Model. The model attained a precision and accuracy measure of 100%, signifying the occurrence of true positives positive predictions, and all classifications were precise. However, the recall rate was 99%, indicating that 1% of actual positive instances might have been missed by the model. Moreover, the Hidden Markov Model outperformed the Support Vector Machines, Decision Trees, and Naive Bayes classifiers. This shows that the Hidden Markov Model's ability to capture sequential dependencies within purchasing behavior led to greater predictive capabilities in comparison to the existing prediction algorithms.

This study looked to understand cardholder purchasing patterns through the application of hidden Markov model, adopting a dynamic approach of identifying most dominant cardholder intentions through transactional transitions across merchant sectors. Through these shifts, merchants and acquirers can tailor their propositions based on established cardholder preferences as informed by their purchasing patterns, enhancing cardholder retention levels and loyalty. Overall, the research showcased a detailed framework that brings together machine learning modelling techniques and practical applications in cardholder purchasing pattern analysis in line with the need for strategy optimization in this space.

VI. Conclusion and Future Work

This research has proved the hidden Markov model's ability to dynamically model cardholder purchasing pattern prediction. The paper has also reviewed the existing algorithms for predicting cardholder purchasing patterns, in particular, support vector machines, Naïve Bayes Classifier and Decision Tree Classifier. By applying expectation maximization, this paper has proved beyond reasonable doubt that hidden Markov model achieves best performance through three key evaluation metrics, i.e., accuracy measure, precision and recall rates. Further, this study has demystified the myth of cardholder purchasing pattern being just about e-commerce credit card fraud detection but a holistic card acceptance ecosystem question, covering all eligible card types, i.e., credit, debit, and prepaid cards transactions on both card present and card not present channels. Additionally, this paper has presented four key pointers contributing to optimal prediction performance, i.e., sector specific modelling through hidden Markov model, distinctive observation state mapping, cardholder profile-based prediction and effective parameter optimization. These key pointers collectively contributed to the accuracy and precision measure of 100%, with a recall rate of 99%, outperforming all other existing models and algorithms in scope.

While our model posted optimal performance across the three metrics, one of the future recommendations would be to slightly normalize the performance for improved practicality by introducing controlled levels of noise into the model, achievable by adding random fluctuations to the emission probabilities, mirroring real-world uncertainties, and making the model more robust to variations in the data.

References

[1] Jurafsky, D., & Martin, J. (2024). Speech and Language Processing with Hidden Markov Models. Standford University. February 3, 2024., pp. 1-17, from https://web.stanford.edu/~jurafsky/slp3/A.pdf.

[2] Bahaa, E., Amnai, M., & Fakhri, Y. (2023). Predicting user behavior using data profiling and hidden Markov model. International Journal of Electrical and Computer Engineering. Vol. 13, No. 5, pp. 5444~5453 2023, ISSN: 2088-8708, DOI: 10.11591/ijece.v13i5.pp5444-5453.

[3] Alif, R., Nazaruddin, Y., & Mandasari, M. (2023). Driver Behavior Prediction Based on Environmental Observation Using Fuzzy Hidden Markov Model. International Journal of Sustainable Transportation Technology. Vol. 6, No. 1, pp. 22-27, from https://unijourn.com/article/64a839c47be3430df6c713cb/5ae99ad07348a8567766abe2/4.html.

[4] Yanrong, W., Wang, C., Wang, H., & Chen, Z. (2024). Prediction of Consumers' Adoption Behaviour of Products with Water Efficiency Labelling Based on Hidden Markov Model. MDPI Journal on Water. Vol 16, pp 1-12. https://doi.org/10.3390/w16010044.

[5] Abdul, K., Owoh, N., Uthmani, O., Ashawa, M., Adejoh, J. & Osamor, J. (2024). Enhancing Credit Card Fraud Detection: An Ensemble Machine Learning Approach. Big Data Cogn. Comput. vol 8, pp 1-7. https://doi.org/10.3390/bdcc8010006.

[6] Saidane, M. (2024). Risk Assessment in Cryptocurrency Portfolios: A Composite Hidden Markov Factor Analysis Framework. Statistics, Journal of Optimization, and Information Computing. Stat. Optim. Inf. Comput. Vol. 12(2), pp 463-487, ISSN 2310-5070, DOI:10.19139/soic-2310-5070-1837.

[7] Xiangzeng, K., Liu, X., Chen, S., Kang, W., Luo, Z., Chen, J., & Wu, T. (2024). Motion Sequence Analysis Using Adaptive Coding with Ensemble Hidden Markov Models. MDPI Journal of Mathematics. Vol 12, pp 1-17, 2024, https://doi.org/10.3390/math12020185.

[8] Hao, D., Huang, J., Liu, Z., L. Li, X. Liu, X. Wang, J., Chen, Wu, Z., & Mao, X. (2024). "Hidden Markov model for spatial analysis of three-dimensional mineralization distribution: Insights into magma flow and mineral exploration targets in the Jinchuan Ni–Cu-(PGE) sulfide deposit, China. Journal of Applied Geochemistry. Vol 162, Pp. 1-6, ISSN 105911, https://doi.org/10.1016/j.apgeochem.2024.105911.

[9] Versfeld, D., & Usman, A. (2024). Principal components-based hidden Markov model for automatic detection of whale vocalisations. Journal of Marine Systems, vol 242, pp 1-23, ISSN 103941, https://doi.org/10.1016/j.jmarsys.2023.103941.

[10] Abukari, D., Daabo, M., & Abdul, A. (2023). A Multi-layered Hidden Markov Model for Real-Time Fraud Detection in Electronic Financial Transactions. Journal of Artificial Intelligence and Data Mining (JAIDM). Vol. 11, No. 4, pp 599-608, DOI:10.22044/jadm.2023.11990.2357.

[11] Unogwu, O., & Filali, Y. (2023). Fraud Detection and Identification in Credit Card Based on Machine Learning Techniques. Wasit Journal of Computer and Mathematics Science. Vol. 2 No. 3 p. 16-22, ISSN: 2788-5879, DOI: https://doi.org/10.31185/wjcm.185.

[12] Mert, O., & Peker, S. (2023). A classification and regression tree algorithm for heart disease modeling and prediction. Healthcare Analytics. Vol 3, pp 1-9, 2023, ISSN 100130, DOI: https://doi.org/10.1016/j.health.2022.100130.

[13] Nabanita, D., Palanisamy, K., Shanmugam, P., & Subramaniam, U. (2023). Life Cycle Cost Analysis of Pumping System through Machine Learning and Hidden Markov Model. MDPI Journal on Processes. Vol 11, 2157, pp 1-23, Article number 2157, DOI: https://doi.org/10.3390/pr11072157.

[14] Luca, B., Wiedenhöft, J., & Schliep, A. (2023). Compressed computations using wavelets for hidden Markov models with continuous observations. PLOS ONE Journals. Vol 18(6), pp1-11, e0286074, https://doi.org/10.1371/journal.pone.0286074.

[15] Abukari, D., Daabo, M., & Barik, A. (2021). Detecting Electronic Banking Fraud on Highly Imbalanced Data using Hidden Markov Models. Earthline Journal of Mathematical Sciences. Vol 7, pp 1-18, E-ISSN : 2581-8147, DOI: https://doi.org/10.34198/ejms.7221.315332.

[16] Cengiz, A., Budak, M., Yagmur, N., & Balcik, F. (2023). Comparison between random forest and support vector machine algorithms for LULC classification. International Journal of Engineering and Geosciences. Vol 8, pp 1-10, e0286074, https://doi.org/10.1371/journal.pone.0286074.

[17] Ashir, J., Dallora, A., Sanmartin, B., Alper, I., Liaqat, A., Hafiz, R., & Anderberg, P. (2023). Early Prediction of Dementia Using Feature Extraction Battery (FEB) and Optimized Support Vector Machine (SVM) for Classification. MDPI Journal on Biomedicines. Vol 11, pp 1-13, Article number 439, DOI: https://doi.org/10.3390/biomedicines11020439.

[18] Shobana, J., Gangadhar, C., Kumar, R., Renjith, P., Bamini, J., & Chincholkar, Y. (2023). E-commerce customer churn prevention using machine learning-based business intelligence strategy. Science Direct Articles on Measurement, Vol 27, pp 1-8, Article number 100728, DOI: https://doi.org/10.1016/j.measen.2023.100728.

[19] Edoardo, R., Voroli, C., & Farcomeni, A. (2023). Quantile-distribution functions and their use for classification, with application to naïve Bayes classifiers. Springer Journal on Statistics and Computing. Vol 33, pp 1-15, article number 55, DOI: https://doi.org/10.1007/s11222-023-10224-4.

[20] Ahmad, A., & Alhameed, R., (2023). Binary Classification of Customer's Online Purchasing Behavior Using Machine Learning. Journal of Techniques. Vol. 5, No. 2, pp 1-24, ISSN 2708-8383, DOI: https://doi.org/10.51173/jt.v5i2.1226.

[21] Norah, A. (2023). An Overview of Credit Card Fraud Detection Learning Techniques. Virginia State University. pp 1-53, Preprint, Article number 11527, from https://easychair.org › preprint_download › tSZp.

[22] Siddhant, S., & Waoo, A. (2023). Decision Tree Machine Learning Approach for Customer Behavior Analysis on Online Product Review. Journal of Emerging Technologies and Innovative Research. Vol. 10, Issue 3, pp 2-6, ISSN 2349-5162 From: www.jetir.org/papers/JETIR2303498.pdf.

[23] Peng, H., Feng, D., Yang, Y., & Wang, Z. (2023). Identifying Subway Commuters Travel Patterns using Traffic Smart Card Data: A Topic Model. School of Civil Engineering, Beijing Jiaotong University. Vol. 3, pp 1-35, Article number 4693938,From https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4693938.

[24] Maryam, M., Karimpour, J., & Mahan, F. (2024). Cyber attacker's next action prediction on dynamic real-time behavior model. Science Direct Articles on Computers and Electrical Engineering, Vol 113, pp 1-20, Article number 109031, DOI: https://doi.org/10.1016/j.compeleceng.2023.109031.

[25] Jandera, A., &Skovranek, T. (2022). Customer Behaviour Hidden Markov Model," MDP Journal, vol 10, pp 1-11, https://doi.org/10.3390/math10081230.

[26] Patel, N., Li, Y., &Hadaegh, A. (2021). Online Transaction Fraud Detection Using Hidden Markov Model and Behaviour Analysis. International Journal of Computer Science and Security. Vol 15, pp 1-14, IJCSS-1610, from https://www.cscjournals.org/library/manuscriptinfo.php?mc=IJCSS-1610.

[27] Nkemnole, E., & Akinsete, A. (2021). Hidden Markov Model using transaction patterns for ATM card fraud detection. Journal of Theoretical and Applied Economics. Vol 29, pp 1-20, Article number 1566, From http://www.ectap.ro/articol.php?id=1566&rid=145.

[28] Anusiuba, I., Overcomer, A., Okechukwu, P., Ekwealor, U., & Anusiuba, A. (2022). The Application of Hidden Markov Model in Credit Card Fraud Detection System. International Journal of Software & Hardware Research in Engineering, vol 10, pp 1-20, ISSN-2347-4890, from https://www.researchgate.net/publication/358661338.

[29] Abdulaziz, A. (2023). A Survey and a Credit Card Fraud Detection and Prevention Model using the Decision Tree Algorithm. Engineering, Technology & Applied Science Research. Vol 13, pp 1-6, Article number 11505, DOI: https://doi.org/10.48084/etasr.6128.

[30] Suhas, P., Srivastav, S., Singh, P., & Sharma, N. (2021). Study of Fraudulent Detection System Using Hidden Markov Model. International Journal of Science Technology & Engineering, Vol 8, pp 1-6, ISSN 2349-784X, From http://www.ijste.org/articles/IJSTEV8I1009.

[31] Thiong'o, F., Mwangi, W., & Oteyo, I (2024). ADUML: A Novel Anomaly Detection Approach for Unsupervised. Preprint submitted to Elsevier. Pp 1-12, Article number 4696208, from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4696208.

[32] Ayesha, A., & Hussain, A (2024). A Performance Analysis of Machine Learning Techniques for Credit Card Fraud Detection. Journal on Artificial Intelligence, Volume 6, pp 1-22, Article Number 047226, DOI: 10.32604/jai.2024.047226.

[33] Swati, S., & Roheet,. B (2021). Process Mining Techniques for Detecting Fraud in Banks: A Study. Turkish Journal of Computer and Mathematics Education, Vol 12, pp 1-18, Article number 358-3375, from: https://turcomat.org/index.php/turkbilmat/article/view/8058/6307.