IJCRT.ORG ISSN : 2320-2882
INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

AN INCEPTION OF BARTER IMMINENT PREDICTION

[1]JENISHA.N, [2]GOWTHAMI.R, [3]VIJAYALAKSHMI.V.R, [4]ANITHA.K
[1, 2, 3] UG Students, [4]Assistant Professor
[1, 2, 3, 4] Electronics and Communication Engineering
[1, 2, 3, 4] PSN College of Engineering and Technology, Tirunelveli, Tamilnadu, India

Abstract: Machine learning is an artificial intelligence (AI) application that provides systems with the ability to automatically learn and improve based on experiences without explicit programming. Machine learning is aimed at developing computer programs that can access data and use it for self-study. The main goal is to give computers the opportunity to learn automatically without human intervention and help, and adjust actions accordingly An Intellect of concept drift poses an additional challenge to existing learning algorithms. Detecting concept changes, such as changing customer preferences for telecommunications services, is very important in terms of forecasting and decision-making applications in a dynamic environment. Specifically, for case-based reasoning systems. This paper presents a novel method for detecting concept drift in a case-based reasoning system. Rather than measuring the actual case distribution, we introduce a new competency model that detects differences through changes incompetence. Eight sets of experiments in three categories show that our method is effective in detecting drift concepts and accurately identifying drift efficiency. These results directly contribute to the research that tackles concept drift in case-based reasoning and competence model studies.

***Index Terms* – Artificial Intelligence, Servlet, System Development Life Cycle, and Data Distribution Management**

I. INTRODUCTION

The concept of drift means the concept of drift means that the statistical nature of the target variables that a model attempts to predict over time in unexpected ways. This is a problem because predictions are becoming less accurate over time. The term concept refers to the amount to be predicted. In a more general sense, it can also be associated with other interesting phenomena, other than target concepts, for example, with input, but in the context of floating concepts, this term usually refers to the target variable, which is the statistical nature of the target variables that are being predicted by Models, change over time in unexpected ways. This is a problem because predictions are becoming less accurate over time. It is a drift of labels Switch time for primarily the same data. So it has a tremendous scope. The proposed system aims to provide a unifying view of the primary and applied concept drift research in data and related areas. Here the future possible outcome of the system is predicted using concept drift methodology. This will make the system to take better action than the existing system which intern improves the production, profit and development of the organization. It analyses the incoming source and available technologies using multiple algorithms and predicts the outcome of the system and provides it to the administrator. In this system, in case if the future prediction fails, the concept drift provides the report based on the failure stating its causes and helps the administrator to analyze and rectify the flaw in the system that made the loss. These algorithms address concept drift from the root sources, which is the distribution drift. Not only can they accurately identify the time of drift, they can also provide the information about the drift.

II PURPOSE OF THE SYSTEM

Our proposed system, the organization can predict its future using the concept drift methodology through which it gets the idea for implementation of the required input. The possibility of the organization or company’s loss could be less and probability of profit will also increase as the future is predicted previously. The efficiency, planning, decisions activity of the organization can be improved and analyzed drastically. It can be used in multiple domains for the best output of the system

III LITERATURE SURVEY

OPTIMIZING REGRESSION MODELS FOR DATA STREAMS WITH MISSING VALUES (Indre zliobaite, Jaakko hollmen-2014) Automated data acquisition systems, such as wireless sensor networks, surveillance systems, or any system that records data in operating logs, are becoming increasingly common, and provide opportunities for making decision on data in real time. Data is generated continuously resulting in a stream of data, and predictive models need to be built and updated online with the incoming data. In addition, the predictive models need to be able to output predictions continuously, and without delays. Automated data acquisition systems are prone to occasional failures. As a result, missing values may often occur. In this paper, we theoretically analyze effects of missing values to the accuracy of linear predictive models. We derive the optimal least squares solution that minimizes the expected mean squared error given an expected rate of missing values. Based on this theoretically optimal solution we propose a recursive algorithm for producing and updating linear regression online, without accessing historical data. Our experimental evaluation on eight benchmark data sets and a case study in environmental monitoring with streaming data validate the theoretical results and confirm the effectiveness of the proposed strategy.

ADAPTIVE RANDOM FORESTS FOR EVOLVING DATA STREAM CLASSIFICATION (Heitor M.Gomes, Albert Bife-2017) Random forests are currently one of the most used machine learning algorithms in the non-streaming (batch) setting. This preference is attributable to its high learning performance and low demands with respect to input preparation and hyper-parameter tuning. However, in the challenging context of evolving data streams, there is no random forests algorithm that can be considered state-of-the-art in comparison to bagging and boosting based algorithms. In this work, we present the adaptive random forest (ARF) algorithm for classification of evolving data streams. In contrast to previous attempts of replicating random forests for data stream learning, ARF includes an effective re sampling method and adaptive operators that can cope with different types of concept drifts without complex optimization for different data sets. We present experiments with a parallel implementation of ARF which has no degradation in terms of classification performance in comparison to a serial implementation, since trees and adaptive operators are independent from one another. Finally, we compare ARF with state-of-the-art algorithms in a traditional test-then-train evaluation and novels delayed la belling evaluation, and show that ARF is accurate and use a feasible amount of resources.

EXTREME LEARNING MACHINE: THEORY AND APPLICATIONS (Gang-Bin Huang-2006) It is clear that the learning speed of feed forward neural networks is in general far slower than required and it has been a major bottleneck in their applications for past decades. Two key reasons behind may be: (1) the slow gradient-based learning algorithms are extensively used to train neural networks, and (2) all the parameters of the networks are tuned iteratively by using such learning algorithms. Unlike these conventional implementations, this paper proposes a new learning algorithm called extreme learning machine (ELM) for single-hidden layer feed forward neural networks (SLFNs) which randomly chooses hidden nodes and analytically determines the output weights of SLFNs. In theory, this algorithm tends to provide good generalization performance at extremely fast learning speed. The experimental results based on a few artificial and real benchmark function approximation and classification problems including very large complex applications show that the new algorithm can produce good generalization performance in most cases and can learn thousands of times faster than conventional popular learning algorithms for feed forward neural networks.

STRUCTURAL PROPERTY-AWARE MULTI-LAYER NETWORK EMBEDDING FOR LATENT FACTOR ANALYSIS (Jie Lu-2018) Multi layer network is a structure commonly used to describe and model the complex interaction between sets of entities/nodes. Network embedding, which aims to project the nodes in the network into a relatively low dimensional space for latent factor analysis, has recently emerged as an effective method for a variety of network-based tasks, such as collaborative filtering and link prediction. However, existing studies of network embedding both focus on the single-layer network and overlook the structural properties of the network. In this paper, we propose four multi layer network embedding algorithms based on Non negative Matrix Factorization (NMF) with consideration given to four structural properties: whole network (NNMF), community (CNMF), degree distribution (DNMF), and max spanning tree (TNMF). Experiments on synthetic data show that the proposed algorithms are able to preserve the desired structural properties as designed. Experiments on real-world data show that multi layer network embedding improves the accuracy of document clustering and recommendation, and the four embedding algorithms corresponding to the four structural properties demonstrate the differences in performance on these two tasks. These results can be directly used in document clustering and recommendation systems.

PCLASS: AN EFFECTIVE CLASSIFIER FOR STREAMING EXAMPLES (Mahardhika pratama-2015) In this paper, a novel evolving fuzzy-rule-based classifier, termed parsimonious classifier (pClass), is proposed. P Class can drive its learning engine from scratch with an empty rule base or initially trained fuzzy models. It adopts an open structure and plug and play concept where automatic knowledge building, rule-based simplification, knowledge recall mechanism, and soft feature reduction can be carried out on the fl-fly with limited expert knowledge and without prior assumptions to underlying data distribution. In this paper, three state-of-the-art classifier architectures engaging multi-input–multi-output, multi model, and round robin architectures are also critically analyzed. The efficient of the p class has been numerically validated by means of real-world and synthetic streaming data, possessing various concept drifts, noisy learning environments, and dynamic class attributes. In addition, comparative studies with prominent algorithms using comprehensive statistical tests have confirmed that the p Class delivers more superior performance in terms of classification rate, number of fuzzy rules, and number of rule-base parameters.

SEMI-SUPERVISED LEARNING IN NON-STATIONARY ENVIRONMENTS GREGORY (Ditzler and Robi Polikar-2011) Learning in non stationary environments, also called learning concept drift concept drift algorithms are trained to detect and track the drifting concepts. While concept drift itself is a significantly more complex problem than the traditional machine learning paradigm of data coming from a fixed distribution, the problem is further complicated when obtaining labeled data is expensive, and training must rely, in part, on unlabeled data. Independently from concept drift research, semi-supervised approaches have been developed for learning from (limited) labeled and (abundant) unlabeled data; however, such approaches have been largely absent in concept drift literature. In this contribution, we describe an ensemble of classifiers based approach that takes advantage of both labeled and unlabeled data in addressing concept drift: available labeled data are used to generate classifiers, whose voting weights are determined based on the distances between Gaussian mixture model components trained on both labeled and unlabeled data in a drifting environment.

NEW ENSEMBLE METHODS FOR EVOLVING DATA STREAMS (Albert bifet-2009) advanced analysis of data streams is quickly becoming a key area of data mining research as the number of applications demanding such processing increases. Online mining when such data streams evolve over time, that is when concepts drift or change completely, is becoming one of the core of classifiers have several advantages over single classifier methods: they are easy to scale and parallelize, they can adapt to change quickly by pruning under-performing parts of the ensemble, and they therefore usually also generate more accurate concept descriptions. This paper proposes a new experimental data stream framework for studying concept drift, and two new variants of Bagging: ADWIN Bagging and Adaptive-Size Huffing Tree (ASHT) Bagging. Using the new experimental framework, an evaluation study on synthetic and real-world data sets comprising up to ten million examples shows that the new ensemble methods perform very well compared to several known methods.

SAND: SEMI-SUPERVISED ADAPTIVE NOVEL CLASS DETECTION AND CLASSIFICATION OVER DATA STREAM (Ahsanul haque-2013) Most approaches to classifying data streams either divide the stream into fixed-size chunks or use gradual forgetting. Due to evolving nature of data streams, finding a proper size or choosing a forgetting rate without prior knowledge about time-scale of change is not a trivial task. These approaches hence suffer from a trade-off between performance and sensitivity. Existing dynamic sliding window based approaches address this problem by tracking changes in classifier error rate, but are supervised in nature. We propose an efficient semi-supervised framework in this paper which uses change detection on classifier confidence to detect concept drifts, and to determine chunk boundaries dynamically. It also addresses concept evolution problem by detecting outlines having strong cohesion among them. Experiment results on bench mark and synthetic data sets show effectiveness of the proposed approach.

IV PROPOSED SYSTEM

The proposed system aims to provide a unifying view of the primary and applied concept drift research in data and related areas. Here the future possible outcome of the system is predicted using concept drift methodology. This will make the system to take better action than the existing system which intern improves the production, profit and development of the organization. It analyses the incoming source and available technologies using multiple algorithms and predicts the outcome of the system and provides it to the administrator. In this system, in case if the future prediction fails, the concept drift provides the report based on the failure stating its causes and helps the administrator to analyze and rectify the flaw in the system that made the loss. These algorithms address concept drift from the root sources, which is the distribution drift. Not only can they accurately identify the time of drift, they can also provide the information about the drift.

ADVANTAGES OF PROPOSED SYSTEM

- In our proposed system, the organization can predict its future using the concept drift methodology through which it gets the idea for implementation of the required input.

- The possibility of the organizations or company's loss could be less and probability of profit will also increase as the future is predicted previously.

- The efficiency, planning, decisions activity of the organization can be improved and analyzed drastically.

- It can be used in multiple domains for the best output of the system

MODULE DESCRIPTION

- **DATA UPLOAD:** The Product Owner will register the information and upload a product. All information and records are stored in the database.

- **CONCEPT DRIFT:** In predictive analytics and machine learning, the concept drift means that the statistical properties of the target variable, which the model is trying to predict, change over time in unforeseen ways. This causes problems because the predictions become less accurate as time passes. The term concept refers to the quantity to be predicted. More generally, it can also refer to other phenomena of interest besides the target concept, such as an input, but, in the context of concept drift, the term commonly refers to the target variable.

- **ERROR RATE-BASED DRIFT DETECTION:** Error rate-based drift detection algorithms form the largest category of algorithms. These algorithms focus on tracking changes in the online error rate of base classifiers. If an increase or decrease of the error rate is proven to be statistically significant, an upgrade process (drift alarm) will be triggered. When a new data instance becomes available for evaluation, DDM detects whether the overall online error rate within the time window has increased significantly. If the confidence level of the observed error rate change reaches the warning level, DDM starts to build a new learner while using the old learner for predictions. If the change reached the drift level, the old learner will be replaced by the new learner for further prediction tasks. To acquire the online error rate, DDM needs a classifier to make the predictions.

- **TEST STATISTICS CALCULATION:** Test statistics calculation is the measurement of dissimilarity or distance estimation. It quantifies the severity of the drift and forms test statistics for the hypothesis test. It is considered to be the most challenging aspect of concept drift detection. The problem of how to define an accurate and robust dissimilarity measurement is still an open question. A dissimilarity measurement can also be used in clustering evaluation, and to determine the dissimilarity between sample sets.

- **DATA DISTRIBUTION-BASED DRIFT DETECTION:** Data Distribution-based Drift Detection Algorithms of this category use a distance function/metric to quantify the dissimilarity between the distribution of historical data and the new data. If the dissimilarity is proven to be statistically significantly different, the system will trigger a learning model upgrade process. These algorithms address concept drift from the root sources, which is the distribution drift. Not only can they accurately identify the time of drift, but they can also provide location information about the drift.

- **ANALYZE PROFIT OR LOSS:** The admin views the product and analyzes profit or loss based on how many people come to the trade center. If the people view the product and command for the product are good or bad. The admin will send the status of the product to product owner.

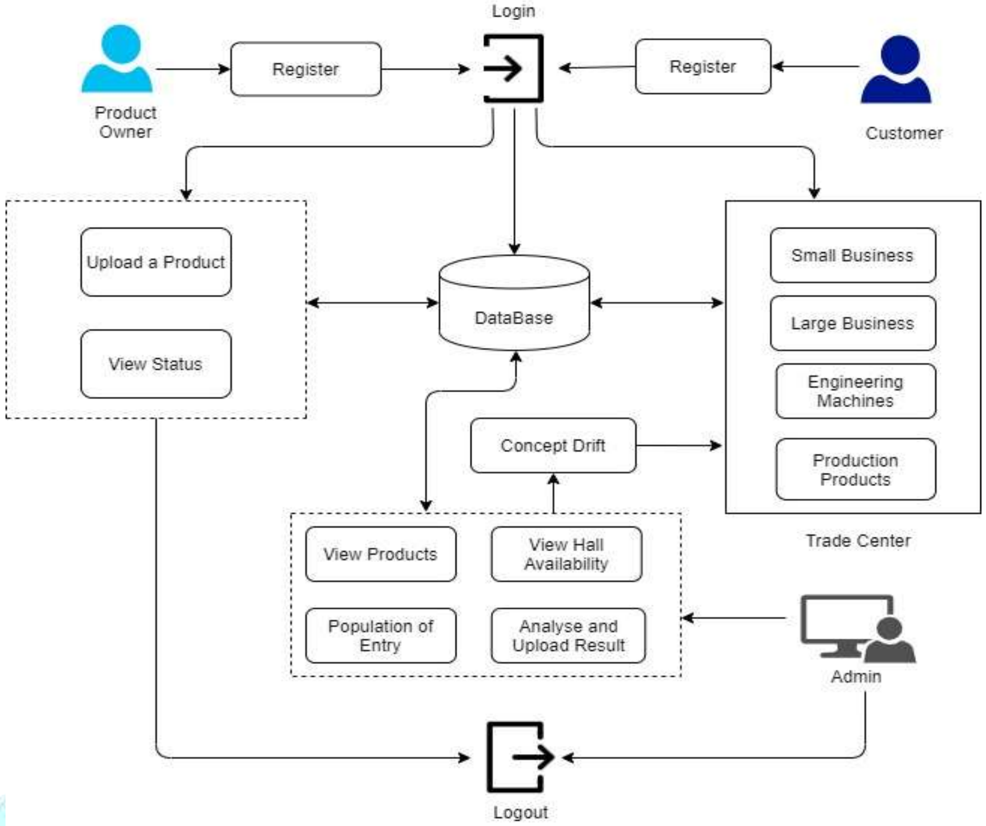

Figure 1: System Architecture of proposed system

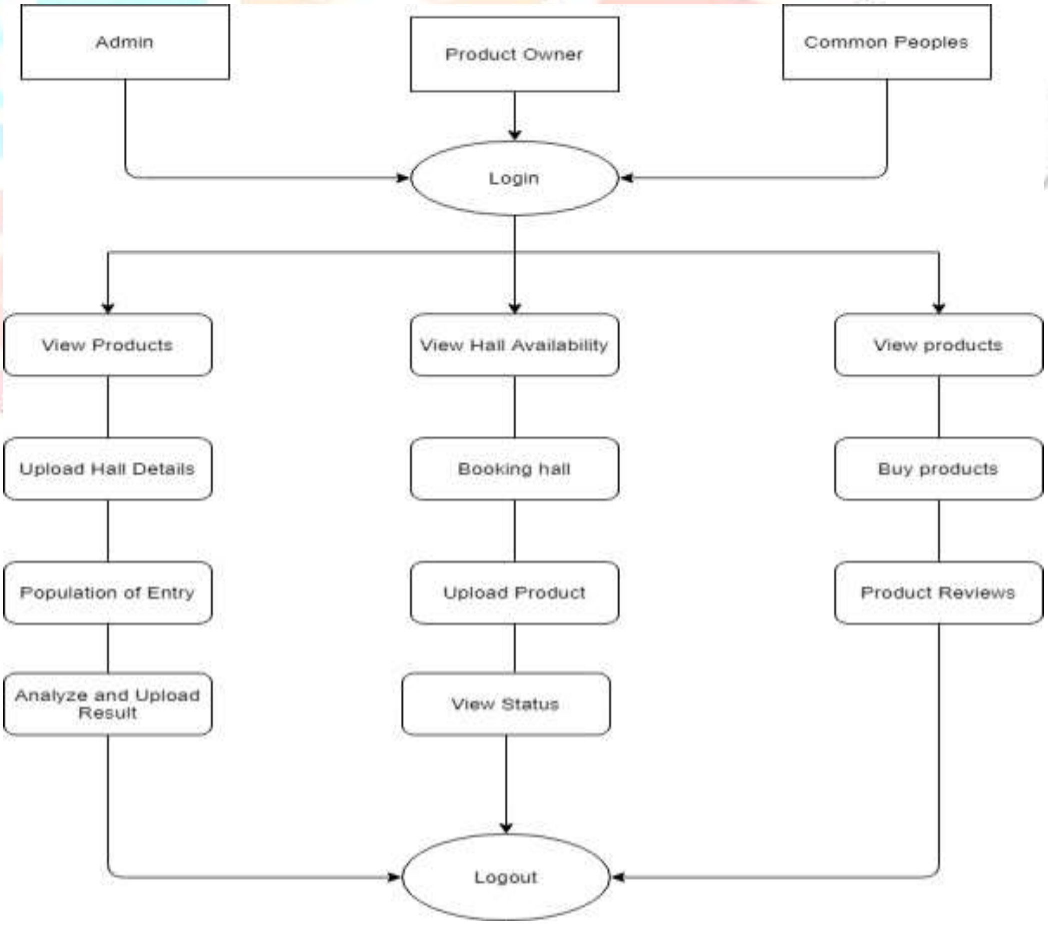

Figure 2: Dataflow diagram 0 level of proposed system

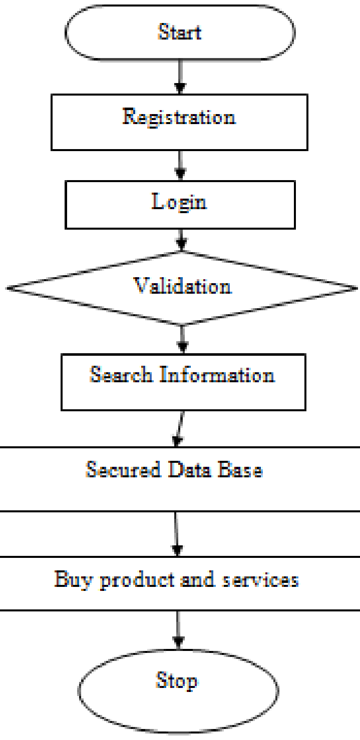

Figure 3: Flowchart of the system

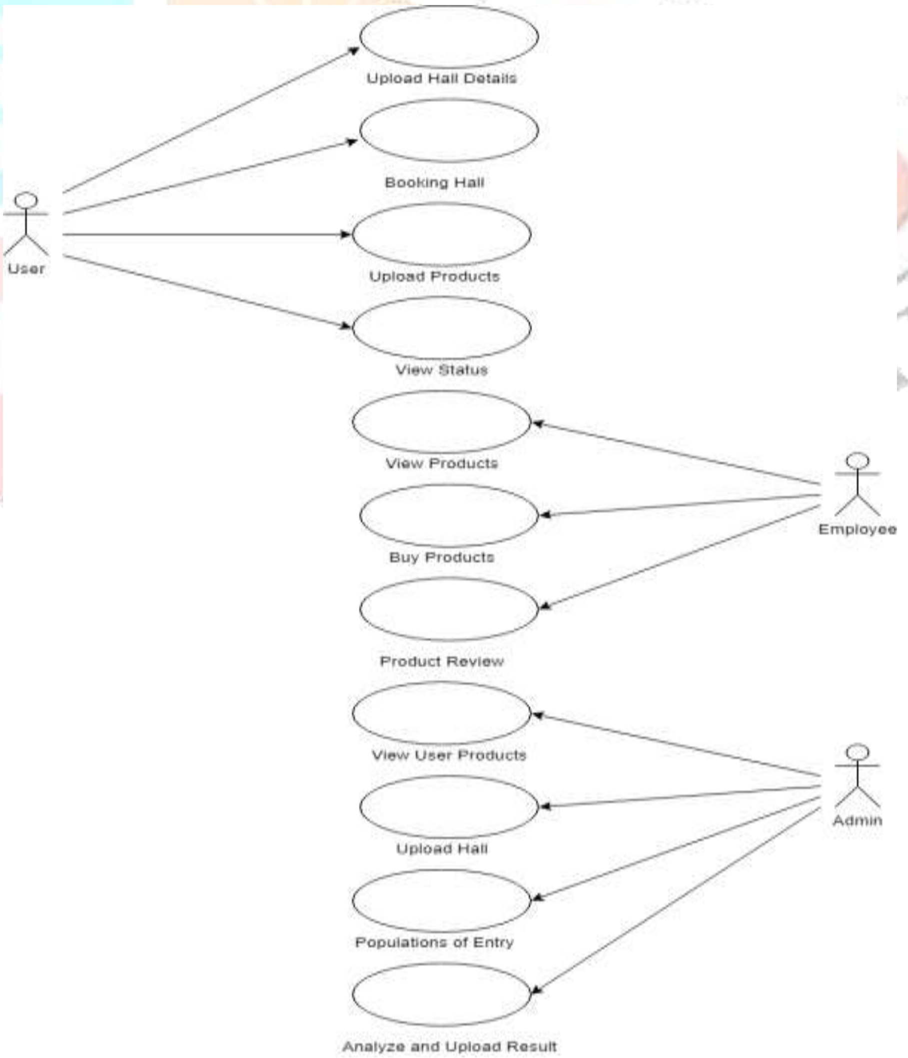

Figure 4: Case Diagram of proposed system

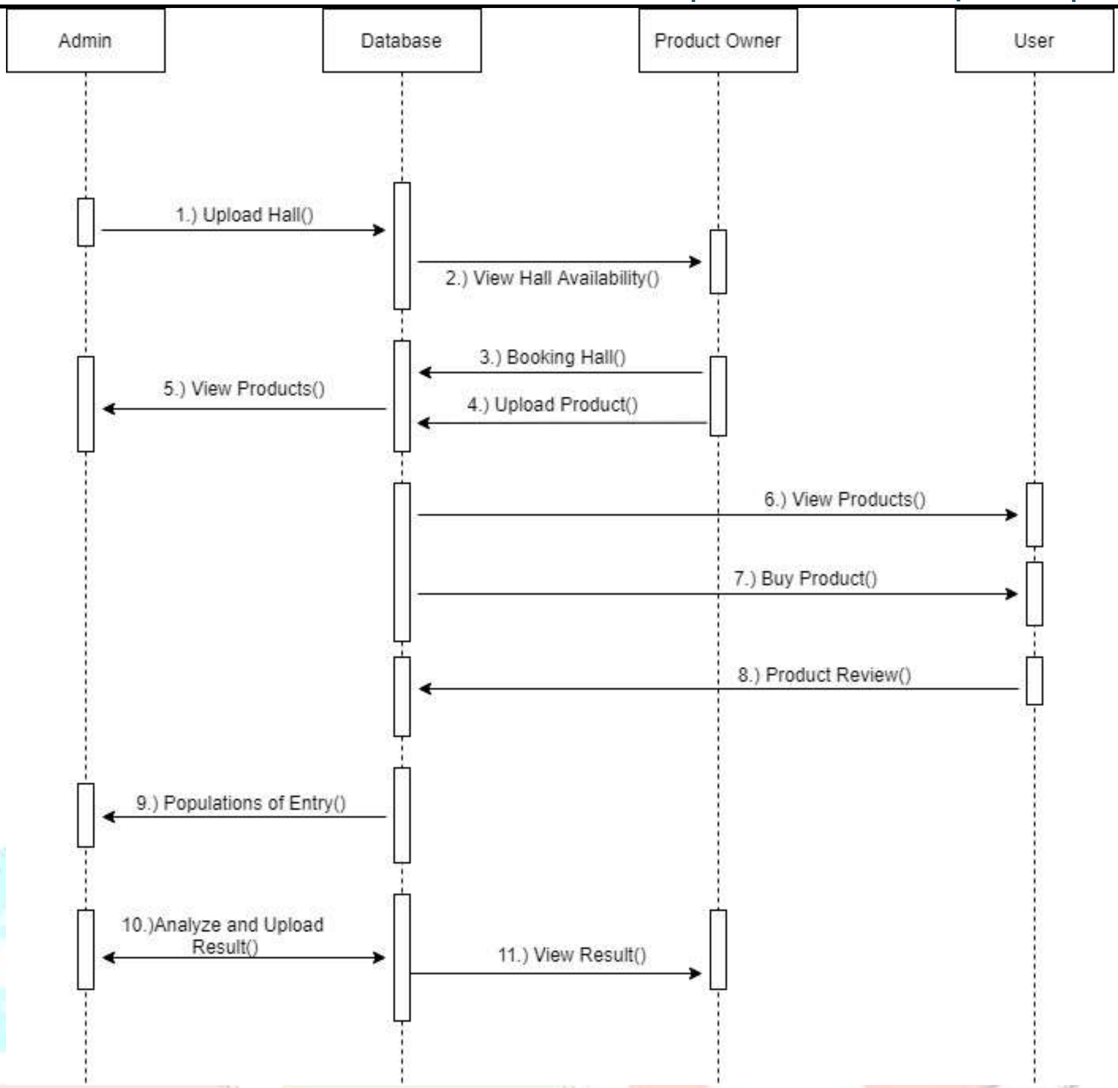

Figure 5: Sequence diagram of proposed system

V RESULT AND DISCUSSION

A Traditional exhibition is a large-scale event that uses a variety of media and means to promote products, corporate image and establish good public relations. They may also want to raise awareness of upcoming products.

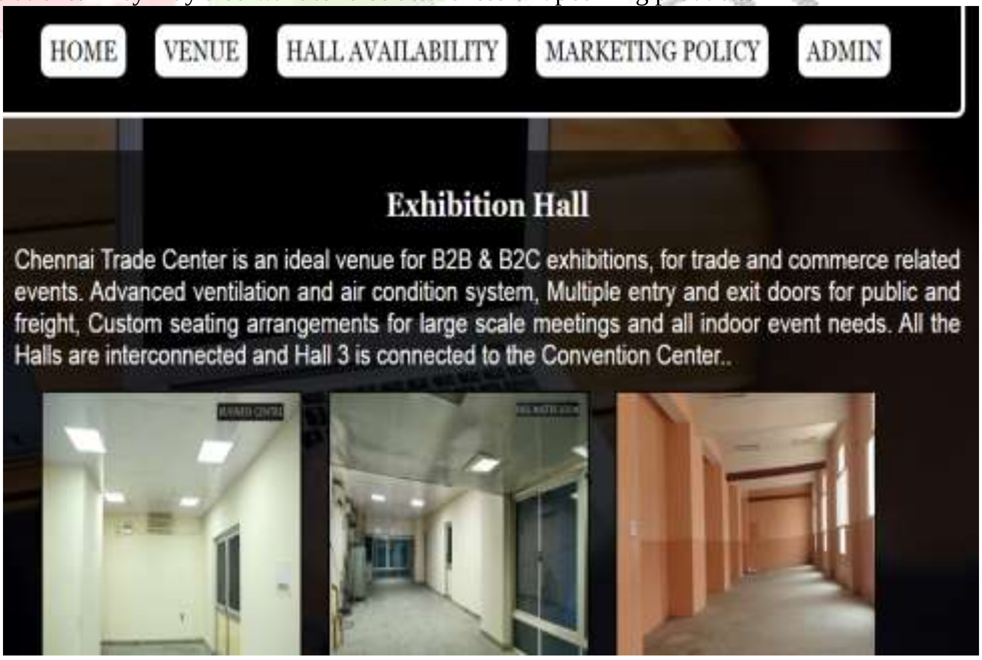

Figure 6: Exhibition Hall

The product owner is a scrum development role for a person who represents the business or user community and is responsible for working with the user group to determine what features will be in the product release. The Product Owner will register the name, contact number, last name, user name, mail id and password finally click the submit button and All information and records are stored in the database.

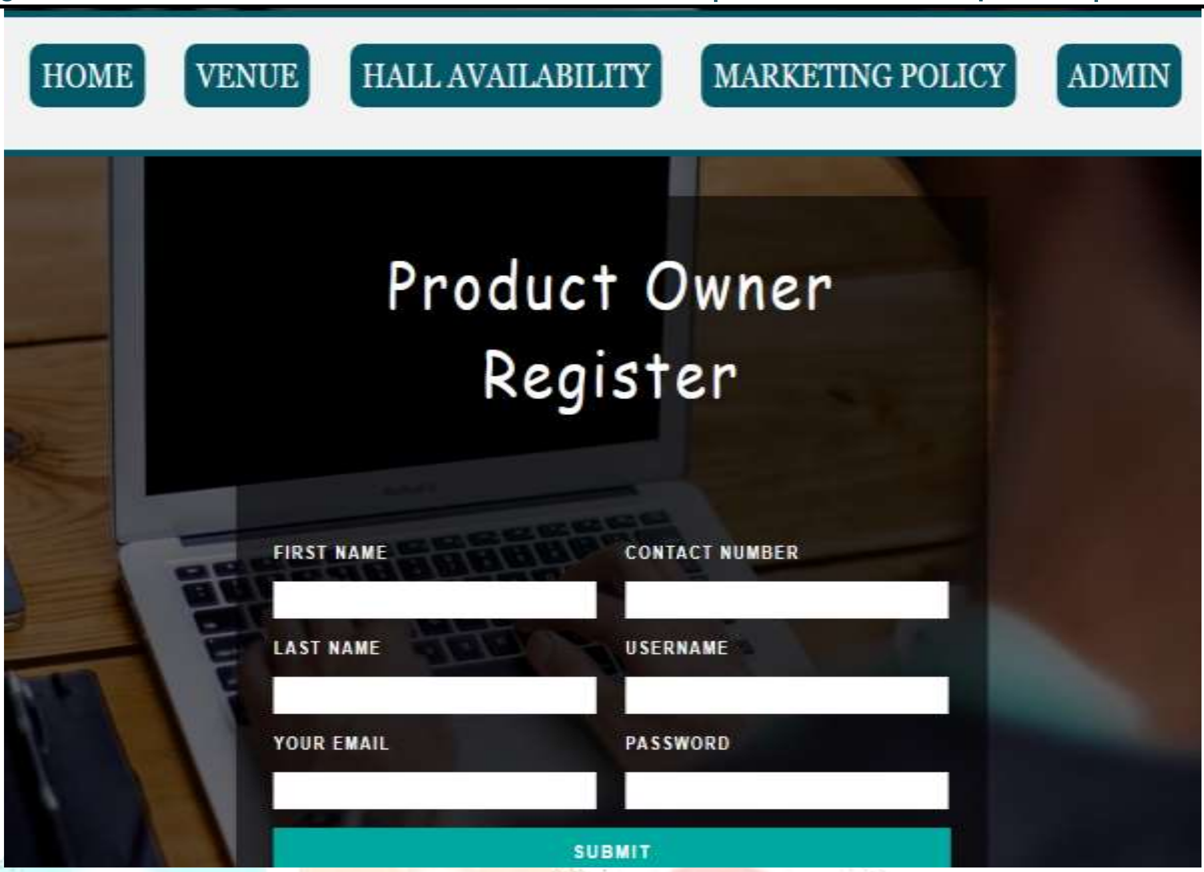

Figure 7: Product owner register

Product owner register hall, hall location, area of the hall, amount rate product name, total investment of the product manufacture, per product price finally upload the product image and All information and records are stored in the database.

Figure 8: BOOK hall

The user will register the name, contact number, last name, user name, mail id and password finally click the submit button and all information and records are stored in the database. An Activity with a UI that allows you to browser settings. Provide a second Activity that allows users to access the share with permission from the administrator. Handle activity lifecycle appropriately.

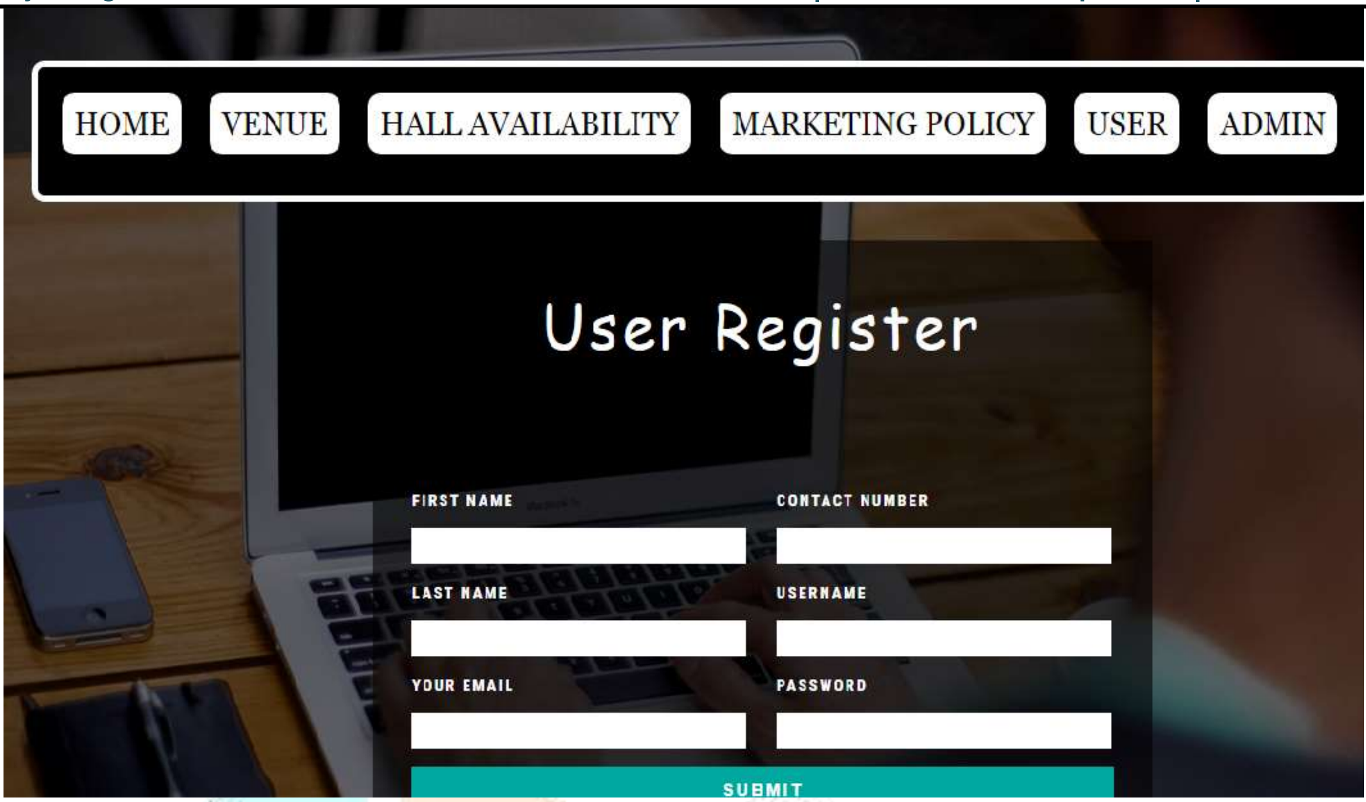

Figure 9: User register

Product page is what define the features, manufacture, uses and lot more, about a certain product. This allows the users to look deeply into what a product offers and how it will benefit them once they buy it. Finally the user enters the review of the product otherwise good or bad etc. The admin will send the status of the product to product owner.

Figure 9 : View product

Today we are going to build a registration system that keeps track of which users are admin and which are normal users. The normal users in our application are not allowed to access admin pages. All users use the same form to login. Admin login the user name and password finally submit the button.

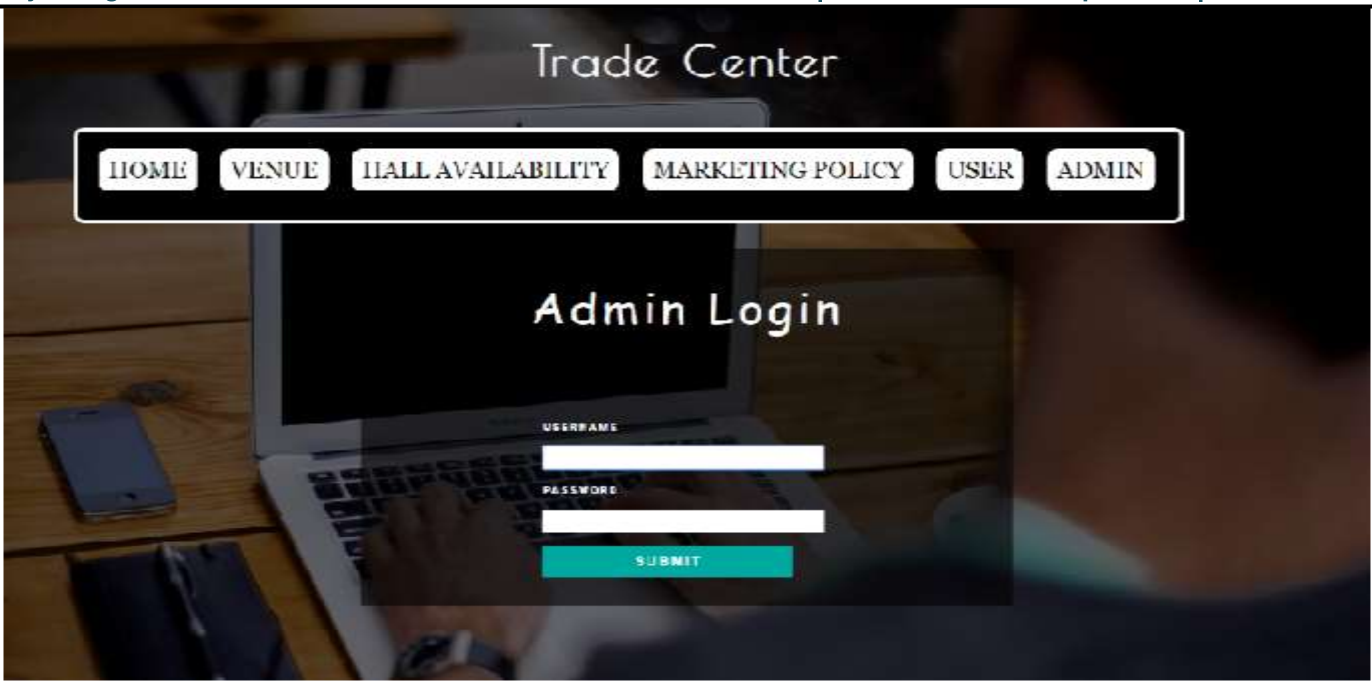

Figure 10: Admin login

This web page displaying product viewing user about a information like name, mail id, and mobile number. How many people's are came to the trade center list. so easy to calculate the viewers.

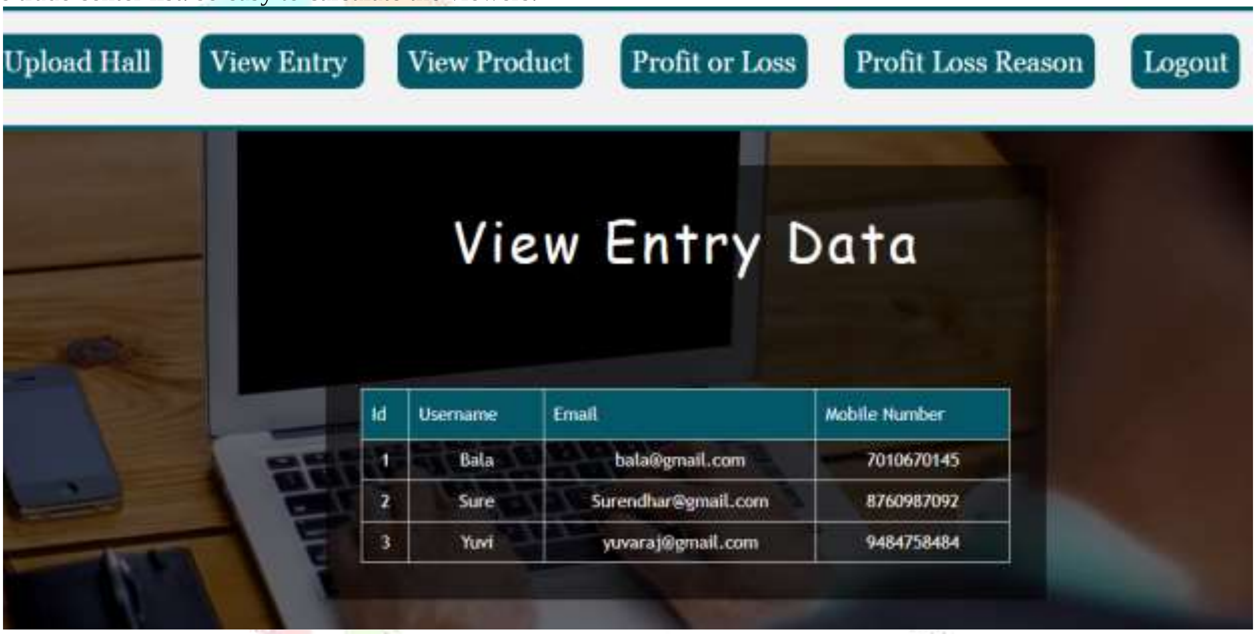

Figure 11: Entry data

The product owner enters the particular trade center location. That particular location product has a profit or loss easy to indentify

Figure 12: Location selection

.The admin views the product and analyzes profit or loss based on how many people come to the trade center. If the people view the product and command for the product are good or bad. The admin will send the status of the product to product owner.

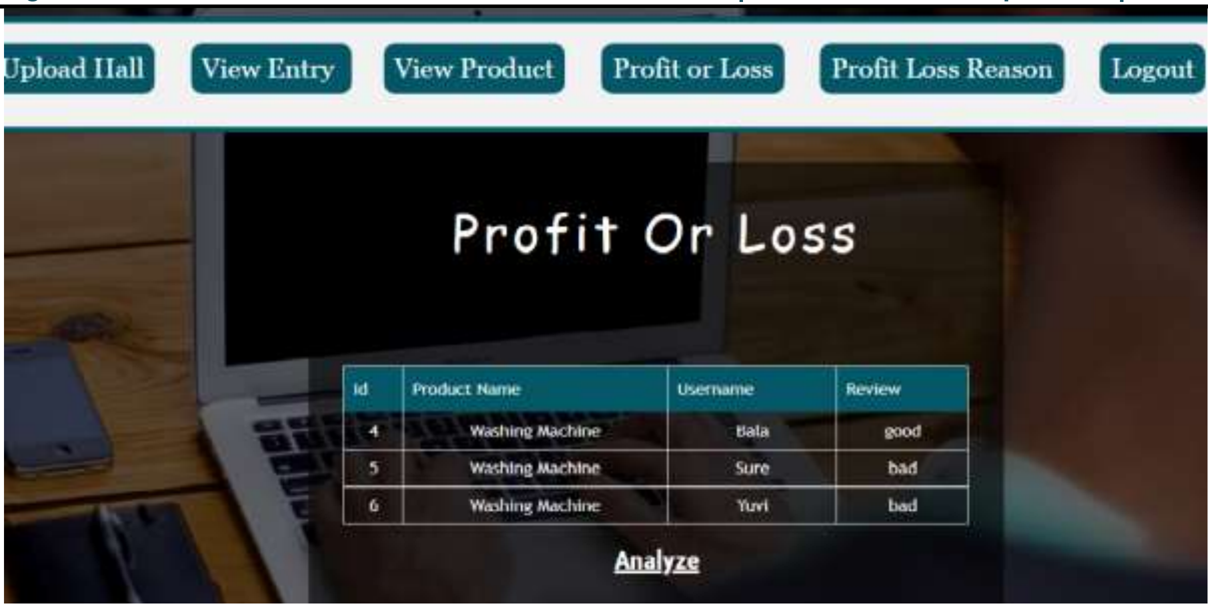

Figure 13: Products and analyze

The admin views the product and analyzes profit or loss based on how many people come to the trade center. If the people view the product and command for the product are good or bad with reason. The admin will send the status of the product to product owner.

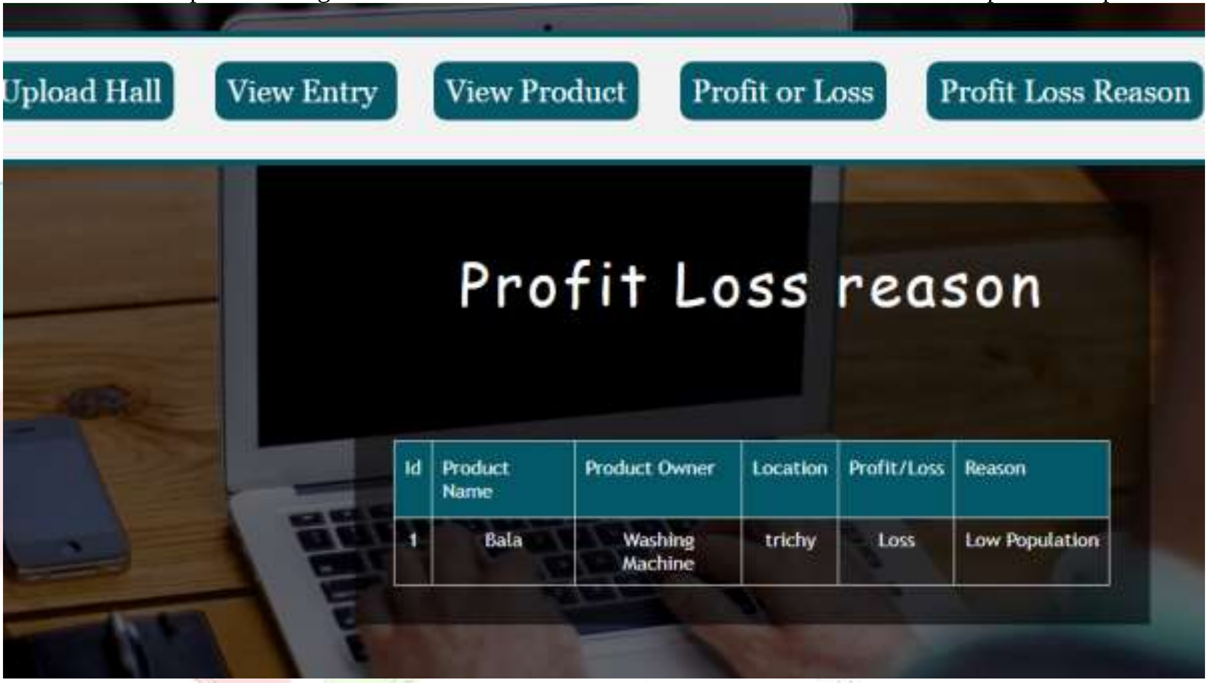

Figure 14: Profit or loss reason

VI CONCLUSION

The future prediction methods are still playing a dominant role in concept drift detection research, while multiple hypothesis test methods emerge in recent years. Regarding to concept drift understanding, all drift detection methods can answer "When", but very few methods have the ability to answer "How" and "Where". Adaptive models and ensemble techniques have played an increasingly important role in recent concept drift adaptation developments. Most existing drift detection and adaptation algorithms assume the ground true label is available after classification/ prediction, or extreme verification latency. Very few research has been conducted to address unsupervised or semi-supervised drift detection and adaptation. Some computational intelligence techniques, such as fuzzy logic, competence model, have been applied in concept drift. There is no comprehensive analysis on real-world data streams from the concept drift aspect, such as the drift occurrence time, the severity of drift, and the drift regions. An increasing number of other research areas have recognized the importance of handling concept drift, especially in big data community.

Concept drift may be present on supervised learning problems where predictions are made and data is collected over time. These are traditionally called online learning problems, given the change expected in the data over time. There are domains where predictions are ordered by time, such as time series forecasting and predictions on streaming data where the problem of concept drift is more likely and should be explicitly tested for and addressed. In the future, the concept drift can be collaborated with top domains in the recent technologies like artificial intelligence, data handling, etc for the data driven applications.

REFERENCES

[1] S.Yu,X.Wang, and J.C.Principe, “Request-and-reverify: Hierarchical hypothesis testing for concept drift detection with expensive labels,” in Proc. 27th Int. Conf. on Artificial Intelligence,2018, pp. 3033–3039

[2] Y.Sun,K.Tang,Z.Zhu, and X.Yao, “Concept drift adaptation by exploiting historical knowledge,” IEEE Trans. Neural Netw.Learning Syst., vol. 29, no. 10, pp. 4822–4832, 2018

[3] C.Alippi,G.Boracchi, and M.Roveri, “Hierarchical change detection test, IEEE Trans. Neural Networks Learn. Syst., vol. 28,no. 2, pp. 246–258, 2017.

[4] S.Yu and Z.Abraham, “Concept drift detection with hierarchical hypothesis testing,” in Proc. 2017 SIAM Int.Conf. Data Mining,2017, pp. 768–776

[5] S.G.Soares and R.Araujo, “An adaptive ensemble of on-line ´extreme learning machines with variable forgetting factor for dynamic system prediction,” Neurocomputing, vol. 171, pp. 693–707, 2016.

[6] D.Liu,Y.Wu, and H.Jiang, “Fp-elm: An online sequential learning algorithm for dealing with concept drift,” Neurocomputing,vol. 207, pp. 322–334, 2016.

[7] D.Han, C.Giraud-Carrier, and S.Li, “Effificient mining of high speed uncertain data streams,” Applied Intelligence, vol. 43, no. 4,pp. 773–785, 2015.

[8] P.M.Gonc¸alves Jr, S.D.de Carvalho Santos,R.S.Barros, and D.C.Vieira, “A comparative study on concept drift detectors,”Expert Systems with Applications, vol. 41, no. 18, pp. 8144–8156,2014.

[9] A. Bifet, G. Holmes, and B. Pfahringer, “Leveraging bagging for evolving data streams,” in Proc. 2010 Joint European Conf. Machine Learning and Knowledge Discovery in Databases, 2010, pp. 135–150.

[10] A.Tsymbal,M.Pechenizkiy,P.Cunningham,and S. Puuronen,“Dynamic integration of classififiers for handling concept drift,”Information Fusion, vol. 9, no. 1, pp. 56–68, 2008.

[11] S.H.Bach and M.Maloof, “Paired learners for concept drift,” in Proc. 8th Int. Conf. Data Mining, 2008,.

[12] N.C.Oza and S.Russell, “Experimental comparisons of online and batch versions of bagging and boosting,” in Proc. 7th ACM SIGKDD Int. Conf Knowledge Discovery and Data Mining, 2001, pp. 359–364.

[13] B.F.J.Manly and D.Mackenzie, “A cumulative sum type of method for environmental monitoring,” Environmetrics, vol. 11,no. 2, pp. 151–166, 2000.

[14] F.Pukelsheim, “The three sigma rule,” The American Statistician,vol. 48, no. 2, pp. 88–91, 1994.