Ch-Oracle A Unified Statistical Framework for Churn Prediction

D. Vignesh[1], Mrs. K. Vasumathi[2], Dr. S. Selvakani[3]

[1]PG Scholar, PG Department of Computer Science, Government Arts and Science College, Arakkonam, Tamil Nadu, India
[2]Assistant Professor, PG Department of Computer Science, Government Arts and Science College, Arakkonam, Tamil Nadu, India
[3]Assistant Professor and Head, PG Department of Computer Science, Government Arts and Science College, Arakkonam, Tamil Nadu, India

Abstract

The User churn stands as a consequential challenge within the realm of online services, posing a substantial threat to the vitality and financial viability of such services. Traditionally, endeavors in churn prediction have transformed the issue into a binary classification task, wherein users are categorized as either churned or non-churned. More recently, a shift towards a more pragmatic approach has been witnessed in the domain of online services, wherein the focus has transitioned from predicting a binary churn label to anticipating the users' return times. This method, aligning more closely with the dynamics of real-world online services, involves the model predicting the specific time of user return at each temporal step, eschewing the simplistic churn label. Nevertheless, antecedent works within this paradigm have grappled with issues of limited generality and imposing computational complexities. This paper introduces ChOracle, an innovative oracle that prognosticates user churn by modeling user return times through the amalgamation of Temporal Point Processes and Recurrent Neural Networks. Furthermore, our approach incorporates latent variables into the proposed recurrent neural network, effectively capturing the latent user loyalty to the system. An efficient approximate variational algorithm, leveraging backpropagation through time, is developed for the purpose of learning parameters within the proposed RNN.

Keywords: framework, customer churn, prediction.

1. Introduction

End-users constitute the primary constituents of any service, be it in the realms of online or offline domains. Consequently, the acquisition and retention of users emerge as pivotal imperatives for service providers.

Recent empirical investigations affirm that the preservation of existing users entails notably lower costs than the procurement of new ones, with the established clientele proving to be more financially advantageous than their nascent counterparts. Consequently, there exists a prevalent inclination to accord heightened consideration to user retention, particularly within the sphere of online services.

Given the escalating prevalence of online services, the phenomenon of user churn, denoting the attrition of clientele, assumes a position of pronounced significance. The intricacies associated with user churn are

exacerbated in online services due to factors such as nominal switching costs, a plethora of competitors, and the ubiquity of complimentary service offerings. Consequently, a considerable body of scholarly endeavors has been directed towards the prognostication of user churn in recent years.
Subsequent to the identification of potential churners, customer relationship management (CRM) systems can strategically engage them through tailored incentives, exemplified by bespoke promotions or gamification methodologies with the overarching objective of perpetuating their allegiance to extant services.
Churn prognostication has been extensively scrutinized across diverse sectors, including the telecommunication industry banking P2P networks online gaming community-based question answering (CQA) services and other virtual platforms. Within the scholarly discourse, diverse conceptualizations of churn are evident, mirroring the nuanced exigencies of distinct service domains.
These definitions, characterizing the phenomenon of churn, may be stratified into three discernible categories. The first delineation pertains to the "Active" category, pertinent to subscription-oriented services. In this context, churn materializes upon the cessation of contractual obligations, concomitant with the departure of the user from the service.
Constituting a substantial contribution, this discourse furnishes a comprehensive survey of paramount predictors of churn within the milieu of subscription services. In consequence, marketing managers are bestowed with discernment regarding the pivotal factors instrumental in the identification of churn. Consequently, the assimilation of this newfound knowledge affords the potential for the refinement and adaptation of marketing strategies.
The predicament confronting the contemporary zenith of expertise lies in the proliferation of myriad systems continually introduced by diverse practitioners and scholars. Not only does this engender a profligate allocation of efforts in repetitively reinventing established methodologies, but it also begets unaddressed nuances within each system, pivotal to its efficacy. Consequently, there ensues a quandary in distilling the authentic requisites and precise specifications inherent in such platforms.
This paper endeavors to expound and elucidate the Churn Management Framework (CMF), an overarching framework that comprehensively addresses all imperatives essential for the development of an efficacious churn management system. The paramount objective is to furnish a tool of heightened utility for the managerial prerogatives of an organization.

2. Related Work

Vapnik [6] Proposed that the Support Vector Machine (SVM) methodology stands as an innovative classification technique grounded in neural network technology and underpinned by statistical learning theory, as expounded by Vapnik in 1995 and 1998. Within the realm of binary classification, SVMs diligently seek to discern a linear optimal hyperplane, wherein the maximization of the margin of demarcation between positive and negative instances becomes the focal objective. This pursuit is tantamount to the resolution of a quadratic optimization problem, wherein the pivotal role is reserved exclusively for the support vectors — namely, the data points situated in close proximity to the optimal hyperplane.
Alok Kumar Rai [2] proposed that the Customer Relationship Management (CRM) endeavors to establish a competitive edge through unparalleled proficiency in comprehending, articulating, delivering, and cultivating extant customer connections. Concurrently it strives to inaugurate and perpetuate novel customer associations. This concept has ascended to prominence as a paramount parlance within

management circles, propelled by the business media and championed by assertive CRM purveyors who espouse it as a universal remedy for the manifold challenges confronting enterprises and managerial cadres. Remarkably, the elucidation of CRM varies markedly among individuals, with some conceptualizing it as synonymous with personalized marketing, while others equate it with the operations of a call center. Furthermore, certain quarters identify database marketing under the umbrella term of CRM, and still, others encapsulate technological solutions within this multifaceted framework.

Tolle's and Meurer [10] proposed that the subsequently, the predictive performance of the aforementioned Support Vector Machine models is benchmarked against Logistic Regression and Random Forests. Our investigation reveals that Support Vector Machines demonstrate commendable generalization prowess when applied to intricate marketing datasets. However, the pivotal role of the parameter optimization process in dictating predictive performance is underscored. Notably, our findings demonstrate that only under the aegis of an optimal parameter selection procedure do Support Vector Machines surpass the traditional Logistic Regression, with Random Forests outperforming both iterations of Support Vector Machines.

Logistic regression stands as a formidable supervised machine learning algorithm tailored for binary classification quandaries, specifically when the target variable assumes a categorical form. Conceptually akin to linear regression, logistic regression distinguishes itself as a classification-oriented counterpart. In essence, logistic regression employs a logistic function, expounded upon by Tolles and Meurer in 2016, to model a binary output variable. The paramount distinction between linear regression and its logistic counterpart lies in the bounded nature of logistic regression's range, confined within the interval [10]. Noteworthy is the fact that logistic regression dispenses with the prerequisite for a linear relationship between input and output variables, owing to the application of a non-linear log transformation to the odds ratio, a concept that will be elucidated in due course.

D. Daley and D. Vere-Jones [4] proposed that the significant constraint inherent in extant studies lies in their proclivity to formulate diverse parametric assumptions regarding the latent dynamics dictating the emergence of observed point patterns. Conversely, in this endeavor, our aim is to proffer a model capable of assimilating a comprehensive and efficacious representation of the underlying dynamics gleaned from event history, eschewing the prerequisite imposition of predetermined parametric structures. The salient advantage of such an approach lies in endowing the proposed model with heightened adaptability to the intricacies of the data, unfettered by fixed parametric constraints. In Section 6, we undertake a comparative analysis between the proposed Recurrent Marked Temporal Point Process (RMTPP) and several alternative processes characterized by specific parametric configurations, thereby substantiating the unparalleled resilience of RMTPP in mitigating the deleterious effects of model misspecification.

Kelleher [8] proposed that the expounds upon the capacity of deep learning to facilitate data-driven decision-making by discerning and extracting intricate patterns from expansive datasets. Its adeptness in assimilating insights from complex data renders deep learning eminently poised to harness the burgeoning expanse of big data and the escalating computational prowess at our disposal. Within the discourse, Kelleher elucidates foundational tenets of deep learning, delineates a historical trajectory of advancements within the discipline, and appraises its contemporary state of the art. The discourse further delves into the pivotal deep learning architectures, encompassing auto encoders, recurrent neural networks, and long short-term networks, while also addressing recent strides such as Generative Adversarial Networks and capsule networks. Additionally, a comprehensive and intelligible exposition on the fundamental algorithms of deep learning—gradient descent and backpropagation—is proffered. Concluding the

narrative, Kelleher contemplates the prospective trajectory of deep learning, surveying major trends, potential evolutions, and formidable challenges on the horizon.
Al-Fawaz [1] proposed that the Journal articles pertaining to this subject predominantly furnish elucidatory expositions on ERP definitions and intricacies, prevalent misconceptions surrounding ERP vis-à-vis business and industrial organizational concerns, diverse vantage points on ERP, empirical investigations appraising industry experiences, contemporary trajectories within the ERP domain, and comprehensive surveys delving into the ERP literature. These inaugural articles proffer illuminative directives tailored for managerial acumen and nascent researchers navigating the labyrinthine landscape of ERPs. The overarching thematic undercurrent accentuates an intimate nexus with Business Process Reengineering (BPR) and a multifaceted spectrum of organizational metamorphoses concomitant with ERP integration. Certain scholarly works endeavor to disentangle the elemental connotations enveloping ERP, engendering retrospectives derived from years of praxis.
David Jacoby [5] proposed that the executive leadership of a prominent consumer products conglomerate cognized that the successful orchestration of a pivotal merger necessitated the seamless integration of its supply chain with that of its newfound strategically Indeed, the substantial economic efficiencies pledged by the merger hinged upon the amalgamation of these two intricate supply chains, both charged with the imperative responsibility of orchestrating the transition of products from their nascent raw material state through the intricate manufacturing process, culminating in the delivery of finished goods into the discerning hands of customers.
Hayes, M.F [7] proposed that the Partner Relationship Management (PRM) serves as a strategic paradigm within the business domain aimed at enhancing the exchange of information between enterprises and their network of channel partners. Web-centric PRM software applications furnish organizations with the capability to tailor and streamline administrative functions by disseminating real-time data, encompassing shipping schedules and other pertinent information, to all partners through the expansive reach of the Internet. Numerous Customer Relationship Management (CRM) providers have incorporated PRM functionalities into their software ecosystems, such as the integration of web-enabled spreadsheets accessible through extranets. The ongoing discourse surrounding PRM revolves around its juxtaposition with Customer Relationship Management (CRM), prompting deliberation on whether the intricate dynamics inherent in channel partnerships necessitate the establishment of PRM as an independent entity or, alternatively, as an integral constituent within the broader framework of CRM.No first line indent for any paragraph except numbered or bulleted paragraphs. Set “Before Text Indent” to the size of approx 3 spaces between text and numbering/bullets for numbered/bulleted paragraphs.
Kotsiantis, S. B [9] proposed that the The Classification algorithm, a form of Supervised Learning, functions as a discerning method employed to ascertain the categorization of novel observations based on prior training data. In the realm of Classification, a computational model assimilates knowledge from a given dataset or set of observations, subsequently assigning new instances to distinct classes or groups . These classes may manifest as binary distinctions, such as Yes or No, 0 or 1, Spam or Not Spam, or more elaborate categories like distinguishing between a cats are a dog. In this context, classes are interchangeable with the terms targets, labels, or categories. Diverging from regression, where the output variable entails a numerical value, Classification exclusively yields categorical outcomes such as "Green or Blue" or "fruit or animal." This distinction underscores the essence of Classification as a supervised learning technique, thereby mandating labeled input data that encapsulates inputs harmonized with their corresponding outputs.

Azadkia, Mona [3] proposed that the K-Nearest Neighbors (KNN) algorithm stands out as a fundamental, albeit crucial, classification algorithm within the expansive field of Machine Learning. Nestled within the domain of supervised learning, KNN finds pervasive application in domains such as pattern recognition, data mining, and intrusion detection. Its ubiquitous relevance in real-world applications is attributable to its non-parametric nature, signifying its abstention from establishing any inherent assumptions regarding the distribution of data. This distinguishes it from other algorithms, such as Gaussian Mixture Models (GMM), which presuppose a Gaussian distribution in the provided dataset. In the realm of KNN, the algorithm operates on a foundation of prior data, often referred to as training data, wherein coordinates are judiciously categorized into groups delineated by a distinctive attribute.

3. Methodology

In Figure1. Explain the configuration of the HDP framework underwent customization to incorporate solely essential tools and systems requisite for traversing all phases of the project at hand. This bespoke amalgamation of installed systems and tools is denominated as the SYTL-BD framework (SyriaTel's big data framework). Within this framework, we integrated the Hadoop Distributed File System (HDFS) for data storage, the Spark execution engine for data processing, Yarn for resource management, Zeppelin as the development user interface, Ambari for system monitoring, and Ranger for system security. Additionally, the Flume System and Scoop tool were utilized to ingest data from external sources into HDFS.

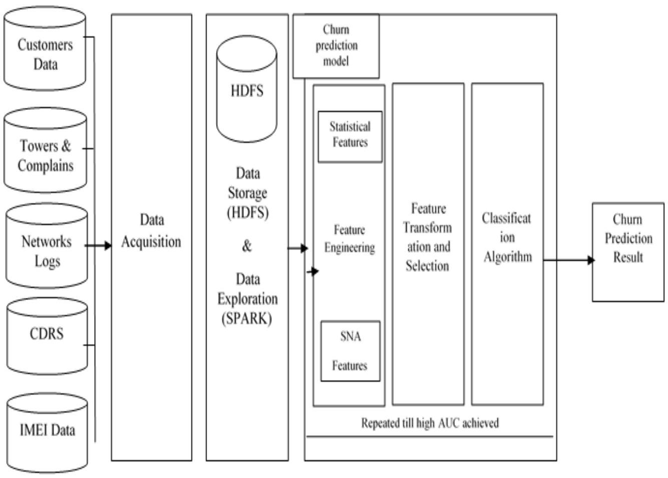

Figure 1. Customer churn prediction

The hardware infrastructure utilized comprised 12 nodes featuring 32 Gigabytes of RAM, 10 Terabytes of storage capacity, and 16 cores per processor for each node. A dataset spanning nine consecutive months was amassed, intended for feature extraction in the churn predictive model.

Spark engine was used in most of the phases of the model like data processing, feature engineering, training and testing the model since it performs the processing on RAM. In addition, there are many other advantages. One of these advantages is that this engine containing a variety of libraries for implementing all stages of machine learning lifecycle.

As expounded in the introductory segment, a contemporary stratagem in the domain of churn prediction involves prognosticating the user's anticipated re-engagement timeframe with the service. The Temporal Point Process (TPP) emerges as a robust mathematical framework, adept at modeling the inherent patterns dictating temporal data dynamics. A predominant constraint within extant investigations employing TPPs to model temporal data lies in their proclivity to impose parametric assumptions concerning the conditional intensity function.

These parameterizations serve as vessels for encapsulating our preconceived knowledge concerning the latent dynamics we endeavor to model. However, in pragmatic scenarios, the veritable model remains elusive. Consequently, diverse specifications for λ

(t) Are explored to refine predictive performance, often culminating in errors attributable to model misjudgment.

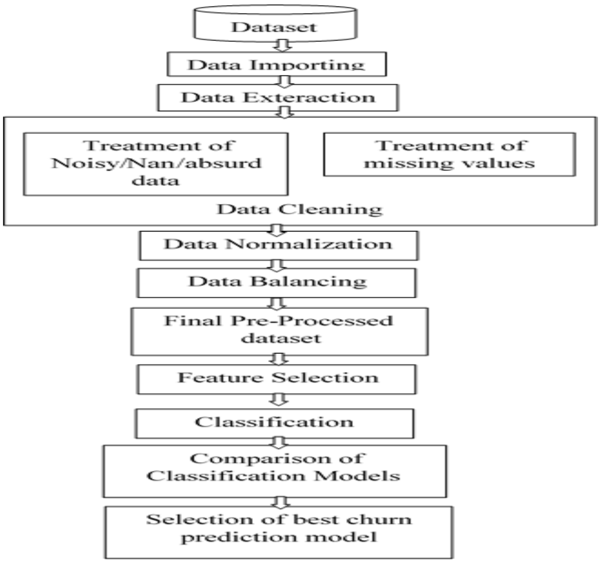

Figure 2. Classification Techniques for Customer Churn Prediction

In Figure 2. Classification Techniques for Customer Churn Prediction. The Dilemma of Model Misspecification. Diverse parameterizations of the conditional intensity function are devised to apprehend and portray distinct forms of historical dependency. The Poisson process, for instance, presupposes a stationarity in duration, while the Hawkes process posits a linear additive impact of past events on the present event. In contrast, the Self-correcting process delineates a non-linear dependency amid past events, and the autoregressive conditional duration model enforces a linear structure upon successive inter-event durations.

Genesis of Markers. Frequently, additional information or covariates, colloquially termed markers, are affiliated with each event. Noteworthy illustrations include the neighborhood-name of a pickup or drop-off location for a NYC taxi, the transactional action of buying or selling in financial transactions, and the diagnosis of a major disease as the marker for a clinical event. Classical temporal point processes can be extended to incorporate marker information through two primary avenues: firstly, by directly integrating the marker into the intensity function; secondly, by considering each marker as an independent dimension, thereby engendering a multidimensional temporal point process.

The former necessitates the specification of a suitable form for the conditional intensity function. However, owing to the augmented complexity induced by markers, a prevailing practice entails the imposition of stringent assumptions asserting the independence of the marker from historical context thereby constricting the model's adaptability. The latter method, in contrast, often contends with the challenge of coping with a profusion of markers, resulting in a sparsity predicament wherein only a sparse subset of events can transpire within each dimension.

A Variational Auto encoder (VAE) introduces a collection of latent random variables, denoted as z, meticulously devised to encapsulate the nuanced variations inherent in the observed variables, x. Operating as an exemplar of a directed graphical model, the joint distribution is articulated as follows:

$$p(x, z) = p(x \mid z)p(z)\ \text{(Equation 3)}.$$

The prior governing the latent random variables, $p(z)$, is commonly stipulated as a Gaussian distribution of simplicity. Meanwhile, the conditional distribution $p(x \mid z)$ assumes the form of an arbitrary observation model, the parameters of which are derived through a parametric function contingent upon z. Significantly, the VAE characteristically parametrizes $p(x \mid z)$ utilizing a highly adaptable function approximator, such as a neural network. While latent random variable models conforming to the structure delineated in Equation are not unprecedented, endowing the conditional $p(x \mid z)$ with the potential for a profoundly non-linear mapping from z to x stands out as a distinctive hallmark of the VAE.

Within this section, we unveil a recurrent iteration of the Variational Auto encoder (VAE), meticulously crafted to proficiently model sequences. Taking cues from less intricate Dynamic Bayesian Networks (DBNs) like Hidden Markov Models (HMMs) and Kalman filters, the conceptualized Variational Recurrent Neural Network (VRNN) overtly delineates the intricacies of interdependence among latent random variables extending across consecutive temporal intervals. In contradistinction to its simpler DBN counterparts, the VRNN not only preserves this capacity for modeling sequential dependencies but also upholds the inherent flexibility to encapsulate profoundly non-linear dynamics.

The Variational Recurrent Neural Network (VRNN) is constituted by an incorporation of a Variational Auto encoder (VAE) at each discrete time step. Notably, these VAEs are conditioned upon the antecedent state variable h_{t-1} of a Recurrent Neural Network (RNN). This augmentation imparts a salutary effect, enabling the VAEs to judiciously incorporate the temporal structure inherent in the sequential data.

Our objective is to substantiate the interconnections among various components of the Churn Management Framework (CMF) and scrutinize the impact of each constituent on the overarching framework. The correlation delineated between the prediction algorithm and the evaluation phase merely elucidates a procedural sequence, rendering the necessity for formal validation superfluous. Furthermore, the bidirectional relationship between the churn index and other customer parameters, such as Customer Lifetime Value (CLV) and Loyalty, has been substantiated by extant research papers.

Regrettably, owing to temporal constraints and the inherent characteristics of the available data, the investigation into the influence of campaign management on data selection lies beyond the purview of this research.

All experimental procedures integral to this research are executed with the support of the Clementine tool, specifically version 11.0 designed for the Windows operating system. Feature selection, the deployment of prediction algorithms, and subsequent evaluation processes are seamlessly orchestrated through the modules available within this software suite. Conversely, the intricacies of data preparation and preprocessing are meticulously managed utilizing the Statistical Package for the Social Sciences (SPSS) version 17.

The experimental endeavors leverage uniform feature selection and prediction algorithms throughout. Each facet of the Churn Management Framework (CMF) undergoes rigorous testing employing distinct methodologies, with the optimal configuration serving as the benchmark for comparative analyses across all experiments. In the validation of individual components within CMF, modifications are introduced to each specific element, and the resultant effects are systematically juxtaposed against the established baseline.

The dataset at our disposal emanates from the Teradata Center at Duke University, encompassing records of seasoned subscribers, specifically those who have maintained their association with a prominent U.S. Telecommunication Company for a minimum duration of six months. Notably, the documented average monthly churn rate within this dataset is purportedly approximately 1.8%. Comprising a total of 172 variables, one serves as an indicator for churn, and the remaining 171 variables are earmarked for predictive purposes. The binary coding of the churn response is manifest in a dummy variable configuration, denoted as churn = 1 should the customer undergo churn, and churn = 0 otherwise.
The determination of potential predictors hinges upon whether the customer severs ties with the company within the 31–60-day interval subsequent to their initial sampling. This temporal parameter of consideration allows practitioners a one-month buffer to strategize and execute measures aimed at customer retention prior to the manifestation of churn. The imposition of a one-month delay in measuring the churn variable is rationalized by the requisite temporal allowance for the effective implementation of proactive customer retention campaigns.
It is noteworthy that a myriad of continuous variables present in the dataset assumes statistical attributes, with some featuring mean values and ranges. Specific continuous variables, exemplified by the range, encapsulate different representations of identical predictors.

4. Experiments and Results

4.1 Software Requirements

- Operating System - Windows 10
- Front End - ASP. Net
- Back End Database - Microsoft SQL Server 2014
- Development Tool - Visual Studio 2017

4.2 Hardware Requirements

- Processor - Dual Core or Above
- Mother Board - Intel Family Motherboard
- RAM - 2 GB and Above
- Hard Disk - 80 GB or Above

4.3 Access Log Extraction

- This module is used to extract the users access log information for the provided web services.
- These data are gathered by storing the user’s login information's and logout information. The system stores the date and time of each login and logout.
- Processing this information the system the average time the users using the services is also calculated.

4.4 Usage Log Extraction

- This module is used to extract the user’s functional usage log information which is complex compare to the previous log extraction process.
- Each functional services and the important pages in the web service runs a background thread which monitors the users clicks and usage time of each page.
- This information is sent to the service provider to calculated usage of each page according to each

user.

4.5 Input Design

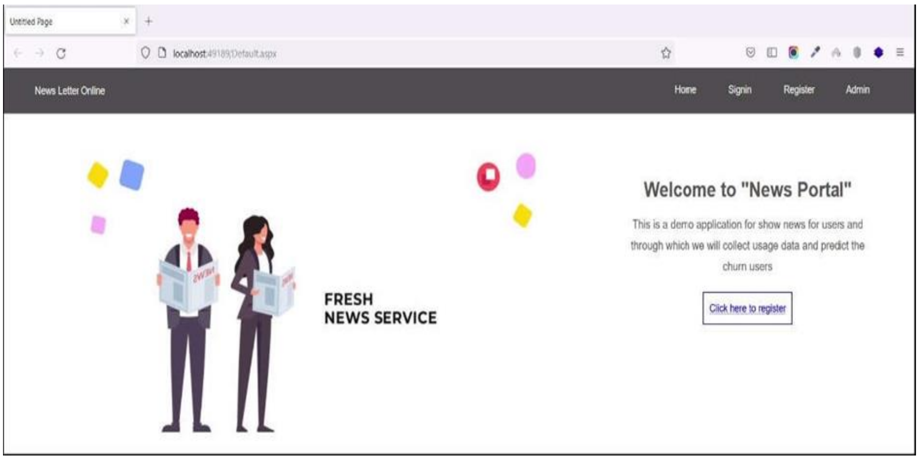

Figure 3. Home page

4.6 User Registration

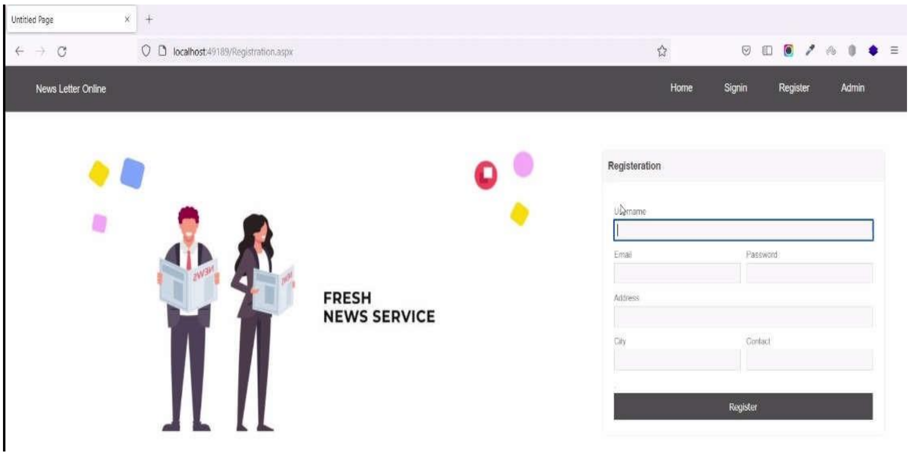

Figure 4. Registration Page

4.7 Login Page

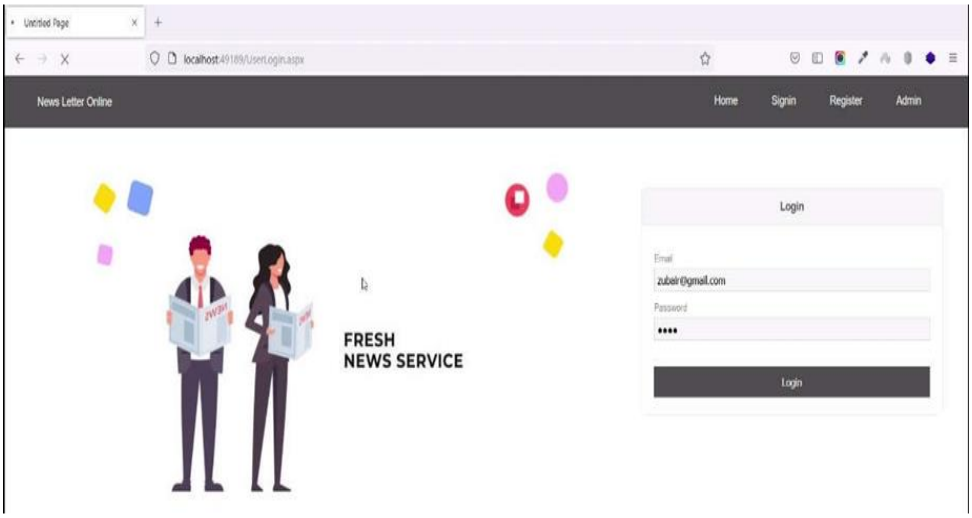

Figure 5. Login Page

4.8 Manual Pattern

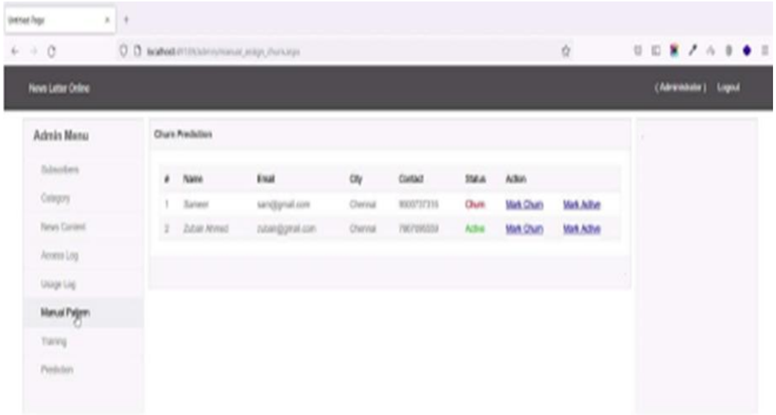

Figure 6. Design for manual pattern

This Manual Pattern is final process of this project. This page says to the customer's status. Name, Email Id, City, Contact, Status and Action. Select the churn customer and normal customer.

4.9 Output Design

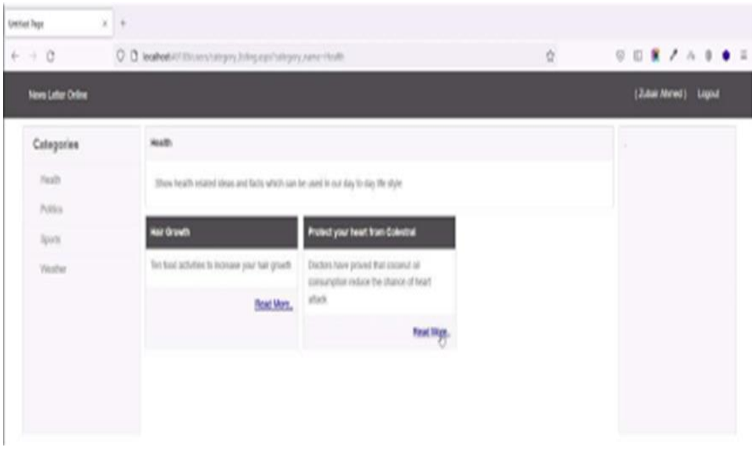

Figure 7. Design for newsletter online

5. Future Work

- It is pertinent to underscore that the initial phase in our churn prediction involves the anticipation of the subsequent lacuna in attendance and the ensuing duration of the next session. Alarms, tailored to the requisites of the application, can then be activated at the subsequent juncture. To illustrate, in the most rudimentary scenario, if the prognosticated values for the upcoming session surpass certain pre-established thresholds, the alarms may be activated—these thresholds being predefined.
- This aligns with the latent conceptualization of churn as expounded in existing literature. Additionally, more intricate thresholds can be contemplated. For instance, these thresholds may be articulated based on the anticipated conduct of the user, encapsulated by expressions such as ^gu_i+1 > E[gu] ^ ^ dui_i+1 > E[du], thereby mirroring the partial delineation of churn. Owing to the application-specific nature of threshold selection, our focus remains on precise value prediction, deferring the meticulous examination of churners and churn alarms to future investigations.
- The IPTV dataset comprises approximately 5000 users and encompasses approximately 1 million

events. When endeavoring to apply the Non-homogeneous Poisson Process (NSR) to the IPTV dataset using a simulation server endowed with a 12GB GPU, we encountered out-of-memory (OOM) errors due to the voluminous data exceeding the available memory capacity. To ameliorate this predicament, we curtailed the prediction timeframe to a mere 500 future hours, resulting in a compromised ability to accurately forecast forthcoming events, as illustrated.

- The Recursive Mean Total Proximity Prediction (RMTPP) method, which solely relies on Recurrent Neural Networks (RNNs) to model event timing, falls short in adequately characterizing the latent patterns governing the temporal dynamics of events. Its performance approximates that of the proposed method solely for the IPTV dataset, distinguished by its extensive training data. However, when confronted with datasets of more limited scope, RMTPP fails to proficiently depict the patterns dictating the temporal dynamics of data.
- Contrastingly, the proposed method, leveraging RNNs to articulate the intensity function of temporal point processes and integrating latent variables into the RNN framework, adeptly captures the latent patterns governing temporal dynamics. Furthermore, it circumvents issues associated with elevated computational complexity.

6. Conclusion

Within this endeavor, we introduced an innovative framework, denominated ChOracle, tailored for prognosticating churn in online services. ChOracle extends temporal point processes to encapsulate the modeling of user absence gaps and session durations. To embody diverse temporal intensities comprehensively, ChOracle leverages recurrent neural networks (RNNs) to articulate the intensity function inherent in temporal point processes. Consequently, the framework exhibits adaptability in modeling various intensity functions. Augmenting the expressive capacity of the model, we introduced latent random variables into the concealed states of the RNN, enabling ChOracle to effectively navigate through highly structured data.

Notably, a Variational lower bound has been derived and adopted as the objective function. The maximization of this objective function through the utilization of backpropagation through time (BPTT) facilitates the comprehensive learning of all parameters. Empirical assessments conducted on real-world datasets underscore the superiority of the proposed ChOracle framework in comparison to state-of-the-art methodologies.

In contemplating future endeavors, one may enhance the predictive efficacy of the proposed method by incorporating more specific data pertaining to user sessions.

References

1. Al-Fawaz, K., Et Al. (2008), ‘Critical Success Factors in Erp Implementation: A Review’, European and Mediterranean Conference on Information Systems.
2. Alok Kumar Rai, Customer Relationship Management Concept & Cases, Prentice Hall Of India Private Limited, New Delhi. 2011.
3. Azadkia, Mona. 2019. “Optimal Choice of K for K-Nearest Neighbor Regression.” E-Print, Arxiv: 1909.05495. Http://Arxiv.Org/Abs/1909.05495.
4. Daley, D.J. And Vere-Jones, D. (1972). A Summary of the Theory of Point Processes. In Lewis (1972), Pp. 299–383.
5. David Jacoby (2009), Guide to Supply Chain Management: How Getting It Right Boosts Corporate

Performance (The Economist Books), Bloomberg Press; 1st Edition, ISBN 978-1576603451.
6. Empirical Inference. Festschrift in Honor of Vladimir N. Vapnik. Schölkopf, Bernhard, Zhiyuan Luo, and Vladimir Vovk Springer Science & Business Media, 2013.
7. Hayes, M.F. And Ref, R. (2003). Partner Relationship Management: The Next Generation of the Extended Enterprise. In Freeland, J.G. (Ed.), The Ultimate Crm Handbook, (Pp. 153-164). Ten Penn Plaza, NY: Mcgraw-Hill.
8. Kelleher, John D, 1974, Material Type: Text Series: The Mit Press Essential Knowledge Series | the Mit Press Essential Knowledge Series Publisher: Cambridge, Massachusetts: The Mit Press, 2019, Description: X, 280 P.: Ill.; 18 Cm ISBN: 9780262537551, Subject(S): Machine Learning Artificial Intelligence.
9. Kotsiantis, S. B. (2007). Supervised Machine Learning: A Review of Classification Techniques. Informatics 31 (2007). Pp. 249 – 268. Retrieved From Ijs Website: Http://Wen.Ijs.Si/Ojs-2.4.3/Index.Php/Informatica/Article/Download/148/140.
10. Tolles J, Meurer Wj. Logistic Regression: Relating Patient Characteristics to Outcomes. Jama. 2016; 316(5):533–4. 10.1001/Jama.2016.7653.