|  |  |  |  |
| --- | --- | --- | --- |
| iid | cite | abstract | |
| 1 | @article{Shinde2022ChurnPA,  title={Churn Prediction and  Prevention for OTT/Tele- Communication},  author={Aniket Shinde},  journal={INTERANTIONAL JOURNAL  OF SCIENTIFIC RESEARCH IN  ENGINEERING AND MANAGEMENT},  year={2022},  url={https://api.semanticscholar  .org/CorpusID:254354035}  } | Customers are becoming more drawn to the quality of service (QoS) offered by businesses in the present. However, the present day shows greater rivalry in offering clients technologically cutting-edge QoS. However, effective customer relationship management systems may help the organization attract new clients, preserve client connections, and enhance client retention by generating more revenue for the company's operations. Churners have always been a major problem for every business that offers services. Churning drives up a company's expenses while also lowering its profit margin. However, it is possible to forecast if a consumer wishes to cancel service using predictive analysis based on historical service usage, service performance, expenditure, and other behavioral patterns. The problem with churn analysis is that it can reveal unnecessary information when used on databases that are combined by a company that owns confidential information. | |
| 2 | @article{Khan2015BehavioralMF,  title={Behavioral Modeling for  Churn Prediction: Early Indicators  and Accurate Predictors of Custom Defection and Loyalty},  author={Muhammad Raza Khan and  Joshua Manoj and Anikate Singh and Joshua Evan Blumenstock},  journal={2015 IEEE International  Congress on Big Data},  year={2015},  pages={677-680},  url={https://api.semanticscholar.  org/CorpusID:2361693}  } | Churn prediction, or the task of identifying customers who are likely to discontinue use of a service, is an important and lucrative concern of firms in many different industries. As these firms collect an increasing amount of large-scale, heterogeneous data on the characteristics and behaviors of customers, new methods become possible for predicting churn. In this paper, we present a unified analytic framework for detecting the early  warning signs of churn, and assigning a “Churn Score” to each customer that indicates the likelihood that the particular individual will churn within a predefined amount of time. This framework employs a brute force approach to feature engineering, then winnows the set of relevant attributes via feature selection, before feeding the final feature-set into a suite of supervised learning algorithms. Using several terabytes of data from a large  mobile phone network, our method identifies several intuitive - and a few surprising - early warning signs of churn, and our best model predicts whether a subscriber will churn with 89.4% accuracy. | |
| 3 | @article{Peng2022ResearchOT,  title={Research on Telecom Customer Churn Prediction Based on GA-XGBoost and SHAP},  author={Ke Peng and Yan Peng},  journal={Journal of Computer and Communications},  year={2022},  url={https://api.semanticscholar .org/CorpusID:254090028}  } | To address the prominent problems faced by customer churn in telecom enterprise management, a telecom customer churn prediction model integrating GA-XGBoost and SHAP is proposed. By using the ADASYN algorithm for data processing on the unbalanced sample set; based on the GA-XGBoost model, the XGBoost algorithm is used to construct the telecom customer churn prediction model, and the hyperparameters of the model are optimized by using the genetic algorithm. The experimental results show that compared with traditional machine learning methods such as GBDT, decision tree, KNN and single XGBoost model, the improved XGBoost model has better performance in recall, F1 value and AUC value; the GA-XGBoost model is integrated with SHAP framework to analyze and explain the important features affecting telecom customer churn, which is more in line with the telecom industry to predict customer the actual situation of churn. | |
| 4 | @inproceedings{Klepac2014DataMM,  title={Data Mining Models as a Tool for Churn Reduction and Custom Product Development in Telecommunication Industries},  author={Goran Klepac},  year={2014},  url={https://api.semanticscholar. org/CorpusID:114286123}  } | This chapter represents the business case in the telecommunication company called Veza, in domain of churn prediction and churn mitigation. The churn project was divided into few stages. Due to limited budget and cost optimization, stage one was concentrated on prospective customer value calculation model based on fuzzy expert system. This helps Veza company to find most valuable telecom subscribers. It also helped company to better understand subscriber portfolio structure. Developed fuzzy expert system also helped Veza company in detection of soft churn. Stage two is profiling and customer segmentation based on time series analysis which provided potential predictors for predictive churn model. The central stage was concentrated on developing traditional predictive churn model based on logistic regression. This calculated probability that subscribers will make churn in next few months. The final stage was dedicated to SNA (Social Network Analysis) model development which found out the most valuable customers from the perspective of existing subscriber network. This model gave the answer that subscribers have the greatest influence on other subscribers in a way what is dangerous if they leave Veza company because they will motivate other subscribers to do the same thing. All three stages made complete churn detection/mitigation solution which take into consideration past behaviour of subscribers, their prospective value, and their strength of influence on other subscribers. This project helped Veza company to decrease churn rate and it gave directions for better understanding customer needs and behaviour which were the base for new product development. | |
| 5 | @article{Barham2023ARO,  title={A Review on Machine Learning-Based Customer Churn Prediction in the Telecom Industry},  author={Sawsan Barham and Nowfal Aweisi and Ala' F. Khalifeh},  journal={2023 9th International Conference on Control, Decision and Information Technologies (CoDIT)},  year={2023},  pages={2659-2664},  url={https://api.semanticscholar. org/CorpusID:264482285}  } | In today's business landscape, companies are facing significant difficulties in achieving positive client interactions and maintaining customers' satisfaction. As a result, businesses are striving to focus on every aspect related to customers and their behaviors to compete in the industry. This has become increasingly important for building customers' loyalties, given the many opportunities available to customize services or products for each customer. To prevent customers from becoming dissatisfied and leaving, businesses are deploying various techniques to efficiently predict customers' behaviors and identify those who may churn or stop using the company's services or products. The rise of machine learning applications has significantly contributed to addressing the challenge of predicting customers' churn rates. Researchers worldwide are now moving towards applying machine learning techniques in this area. This paper aims to present a review of various studies conducted from 2019 to 2022 that utilized machine-learning techniques to predict customers' churn in the telecom industry. The paper summarizes the different machine learning algorithms that are used for customers' churn prediction, with a particular focus on the telecom industry, as well as their accuracy, to provide insights into the effectiveness of these techniques in addressing customers' churn in the telecom industry. | |
| 6 | @article{Cheng2023TowardsDT,  title={Towards data-driven tele-medicine intelligence: community-based mental healthcare paradigm shift for smart aging amid COVID-19 pandemic},  author={Lan Cheng and Wk Chan and Yi Peng and Harry Qin},  journal={Health Information Science and Systems},  year={2023},  volume={11},  url={https://api.semanticscholar. org/CorpusID:257498970}  } | It’s benefcial to involve elderly and gerontechnology stakeholders as part of Community-Based Participatory Research (CBPR) before and throughout the developing and delivery phases an integrated and age-friendlydigital intervention. The challenges in applying and disseminating telemedicine refected by the elderly and caregivers can be used as important input for further development and indicators for the sustainable and integrated elderly  primary care framework. | |
| 7 | @article{Hooda2022AnOK,  title={An Optimized Kernel MSVM Machine Learning-based Model for Churn Analysis},  author={Pankaj Hooda and Pooja Mittal},  journal={International Journal of Advanced Computer Science and Applications},  year={2022},  url={https://api.semanticscholar. org/CorpusID:249297874}  } | Customer churn is considered as a significant issue in any industry due to various services, clients, and commodities. A massive amount of data is being created from e-commerce  services and tools. Analytical data and machine learning-based approaches have been implemented and utilized for CA (churn analysis) to design a plan, i.e., required to comprehend the rationale for the CC (Customer Churn) and to generate a profitable and actual customer holding program. The analytics and machine learning approaches mainly focus on customer profiling, CC classification, and detection of features that affect  churn. However, there are no specific techniques which can be used to determine how often a prospective customer is inclined to cover all the expenses whether they are churned or not. In this paper, an Optimized Kernel MSVM classification model is proposed to predict and classify churn. In the proposed work, MSVM algorithm has been used for classification. The kernel PCA and ALO optimizer method has been used for Feature  extraction and selection. The proposed model Optimized Kernel MSVM has been implemented on Tele-communication sector customer churn database to demonstrate the proposed model's generalization ability. The Optimized Kernel MSVM model has  achieved an accuracy of 91.05%, AUC 85% being maximum and reduced the RMSE score to 2.838. The implementation shows that both churn detection and classification may be examined at the same time while maintaining the highest overall accuracy and AUC. | |
| 8 | @article{Akbas2021ACF,  title={A Computational Framework Towards the Tele-Rehabilitation of Balance Control Skills},  author={Kubra Akbas and Carlotta Mummolo},  journal={Frontiers in Robotics and AI},  year={2021},  volume={8},  url={https://api.semanticscholar. org/CorpusID:235373150}  } | Mobility has been one of the most impacted aspects of human life due to the spread of the COVID-19 pandemic. Home confinement, the lack of access to physical rehabilitation, and prolonged immobilization of COVID-19-positive patients within hospitals are three major factors that affected the mobility of the general population world-wide. Balance is one key indicator to monitor the possible movement disorders that may arise both during the COVID-19 pandemic and in the coming future post-COVID-19. A systematic quantification of the balance performance in the general population is essential for preventing the appearance and progression of certain diseases (e.g., cardiovascular, neurodegenerative, and musculoskeletal), as well as for assessing the therapeutic outcomes of prescribed physical exercises for elderly and pathological patients. The methodology proposed in this research can support the development of innovative technologies for smart and connected home-care solutions for physical therapy rehabilitation. | |
| 9 | @article{Liu2015ResearchMO,  title={Research Model of Churn Prediction Based on Customer Segmentation and Misclassification Cost in the Context of Big Data},  author={Y. Liu and Yongrui Zhuang},  journal={Journal of Computational Chemistry},  year={2015},  volume={03},  pages={87-93},  url={https://api.semanticscholar. org/CorpusID:28281425}  } | Enterprises have vast amounts of customer behavior data in the era of big data. How to take advantage of these data to evaluate custom forfeit risks effectively is a common issue faced by enterprises. Most of traditional customer churn predicting models ignore customer segmentation  and misclassification cost, which reduces the rationality of model. Dealing with these deficiencies, we established a research model of customer churn based on customer segmentation and misclas-  sification cost. We utilized this model to analyze customer behavior data of a telecom company. The results show that this model is better than those models without customer segmentation and misclassification cost in terms of the performance, accuracy and coverage of model | |
| 10 | @article{Yang2018HapticTO,  title={Haptic tele-driving of wheeled mobile robot over the internet via PSPM approach: theory and experiment},  author={Hyunsoo Yang and Zhiyuan Zuo and Dongjun Lee},  journal={Advanced Robotics},  year={2018},  volume={32},  pages={683 - 696},  url={https://api.semanticscholar. org/CorpusID:52077846}  } | We propose novel haptic tele-driving control frameworks of a wheeled mobile robot (WMR) over the imperfect Internet communication network with varying delay and packet loss. We consider both the dynamic and kinematic WMRs and their various tele-driving modes. By utilizing passive set-position modulation framework, we can guarantee two-port passivity or passivity/stability combination of the closed-loop tele-driving system with some theoretical performance measures. Experiments are performed to show the efficacy of the proposed frameworks using the Internet-emulated communication and a custom-built dynamic/kinematic WMR. | |
| 11 | @article{Raghuraman2017AVL,  title={A Visual Latency Estimator for 3D Tele-Immersion},  author={Suraj Raghuraman and K. Bahirat and Balakrishnan Prabhakaran},  journal={Proceedings of the 8th ACM on Multimedia Systems Conference},  year={2017},  url={https://api.semanticscholar. org/CorpusID:28700203}  } | 3D Tele-Immersion systems allow geographically distributed users to interact in a virtual world using their "live" 3D models. The capture, reconstruction, transfer, and rendering of these models introduce significant latency into the system. Implicit Latency (ℒ') can be estimated using system clocks to measure the time after the data was received from the RGB-D camera, till the request to render the result. The Observed Latency (ℒ) between a real world event and the event being rendered on the display, cannot be accurately represented by ℒ' since ℒ' ignores the time taken to capture, or update the display, etc. In this paper, a Visual Pattern based Latency Estimation (VPLE) approach is introduced to calculate the real world visual latency of a system without the need for any custom hardware. VPLE generates a constantly changing pattern that is captured and rendered by the 3DTI system. An external observer records both the pattern and the rendered results at high frame rates. ℒ is estimated by calculating the difference between the generated andrendered patterns. VPLE is extended to allow ℒ estimation between geographically distributed sites. Evaluations show that the accuracy of VPLE depends on the refresh rate of the pattern, and is within 4ms. ℒ of a distributed 3DTI system implemented on the GPU is significantly lower than the CPU implementation, and is comparable to video streaming. It is also shown that the ℒ' estimates for GPU based 3DTI implementations are off by almost 100% compared to the ℒ | |
| 12 | @article{Enayati2017SkillbasedHC,  title={Skill-based human–robot cooperation in tele-operated path tracking},  author={Nima Enayati and Giancarlo Ferrigno and Elena De Momi},  journal={Autonomous Robots},  year={2017},  volume={42},  pages={997 - 1009},  url={https://api.semanticscholar. org/CorpusID:25112243}  } | This work proposes a shared-control tele-operation framework that adapts its cooperative properties to the estimated skill level of the operator. It is hypothesized that different aspects of an operator’s performance in executing a tele-operated path tracking task can be assessed through conventional machine learning methods using motion-based and task-related features. To identify performance measures that capture motor skills linked to the studied task, an experiment is conducted where users new to tele-operation, practice towards motor skill proficiency in 7 training sessions. A set of classifiers are then learned from the acquired data and selected features, which can generate a skill profile that comprises estimations of user’s various competences. Skill profiles are exploited to modify the behavior of the assistive robotic system accordingly with the objective of enhancing user experience by preventing unnecessary restriction for skilled users. A second experiment is implemented in which novice and expert users execute the path tracking on different pathways while being assisted by the robot according to their estimated skill profiles. Results validate the skill estimation method and hint at feasibility of shared-control customization in tele-operated path tracking. | |
| 13 | @inproceedings{Fink2016SmartOT,  title={Smart ophthalmics: the future in tele-ophthalmology has arrived},  author={Wolfgang Fink and Mark A. Tarbell and Kevin Garcia},  booktitle={Defense + Security},  year={2016},  url={https://api.semanticscholar .org/CorpusID:31400218}  } | Smart Ophthalmics© extends ophthalmic healthcare to people who operate/live in austere environments (e.g., military, third world, natural disaster), or are geographically dispersed (e.g., rural populations), where time, cost, and the possibility of travel/transportation make access to even adequate medical care difficult, if at all possible. Operators attach optical devices that act as ophthalmic examination extensions to smartphones and run custom apps to perform examinations of specific areas of the eye. The smartphone apps submit over wireless networks the collected examination data to a smart remote expert system, which provides in-depth medical analyses that are sent back in near real-time to the operators for subsequent triage. | |
| 14 | @article{Yu2016ParticleCO,  title={Particle classification optimization-based BP network for telecommunication customer churn prediction},  author={Ruiyun Yu and Xuanmiao An and Bo Jin and Jiashun Shi and Oguti Ann Move and Yonghe Liu},  journal={Neural Computing and Applications},  year={2016},  volume={29},  pages={707 - 720},  url={https://api.semanticscholar. org/CorpusID:9237695}  } | Customer churn prediction is critical for telecommunication companies to retain users and provide customized services. In this paper, a particle classification optimization-based BP network for telecommunication customer churn prediction (PBCCP) algorithm is proposed, which iteratively executes the particle classification optimization (PCO) and the particle fitness calculation (PFC). PCO classifies the particles into three categories according to their fitness values, and updates the velocity of different category particles using distinct equations. PFC calculates the fitness value of a particle in each forward training process of a BP neural network. PBCCP optimizes the initial weights and thresholds of the BP neural network, and brings remarkable improvement on customer churn prediction accuracy. | |
| 15 | @article{Tsai2019AnIC,  title={An Intelligent Customer Churn Prediction and Response Framework},  author={Tien-Yu Tsai and Chin-Teng Lin and Mukesh Prasad},  journal={2019 IEEE 14th International Conference on Intelligent Systems and Knowledge Engineering (ISKE)},  year={2019},  pages={928-935},  url={https://api.semanticscholar .org/CorpusID:221281774}  } | Customer retention is one of the most important issues for companies. Companies always seek to reduce customer churn in order to increase the customer lifetime value and reduce the cost of acquisition of new customers. By focusing on customer churn prediction and identification, companies can predict in advance which customers are going to churn and therefore decrease customers churn rate through related personalized actions. The key issue here is how to predict customer churn at an early stage. This paper identifies related issues in customer churn prediction and provides new definitions and classifications on customer churn identification and strategies. This paper also establishes a customer churn prediction and response framework consists of three main stages: customer churn prediction, customer churn understanding and customer churn response. The framework presents the characteristics and challenges of related stages of customer churn as well. These outcomes can be used for customized or personalized product and service developments, to improve customer service efficiency and related decision-making more effective and more particularly enabling strategic promotion campaigns to customers with high churn risk. | |
| 16 | @inproceedings{Proczuk2018ChurnRI,  title={Churn Risk Identification as an Important Aspect of Marketing Controlling – the Case of a German Start-Up Company},  author={Anna Prończuk},  year={2018},  url={https://api.semanticscholar. org/CorpusID:169325510}  } | The purpose of this paper is to understand possible methods of identify-  ing churn risk in small and medium-sized start-up companies.  Design/methodology/approach – This paper describes the case study of a German IT  start-up company and its churn risk identification approach. All presented insights are  based on the company’s internal documentation. Additionally, the author conducted  online surveys addressing 50 client teams asking them to assess the occurrence probabil-  ity of the most common risk. The surveys have been conducted every month (August-  -November 2017) with 16 middle and upper managers | |
| 17 | @inproceedings{Alugubelli2018EnhancingSB,  title={Enhancing Subscription Based Business by Predicting Churn Likelihood},  author={Sujal Reddy Alugubelli and Smitha Etlapur and Mounika and Kondamudi},  year={2018},  url={https://api.semanticscholar. org/CorpusID:51790461}  } | Customer retention is a challenge faced by most businesses in today’s competitive market. Predicting customer churn would help a subscription business such as KKBox in creating substantial difference in their revenue stream. This paper describes work relating to predicting churn likelihood using SAS® 9.4, SAS® Enterprise Miner for data cleaning, preparation and analysis. | |
| 18 | @inproceedings{Can2017ChurnPF,  title={Churn Prediction for Mobile Prepaid Subscribers},  author={Zehra Can and Erinç Albey},  booktitle={International Conference on Data Technologies and Applications},  year={2017},  url={https://api.semanticscholar. org/CorpusID:4368728}  } | In telecommunication, mobile operators prefer to acquire postpaid subscribers and increase their incoming revenue based on the usage of postpaid lines. However, subscribers tend to buy and use prepaid mobile lines because of the simplicity of the usage, and due to higher control over the cost of the line compared to postpaid lines. Moreover the prepaid lines have less paper work between the operator and subscriber. The mobile subscriber can end their contract, whenever they want, without making any contact with the operator. After reaching the end of the defined period, the subscriber will disappear, which is defined as “involuntary churn”. In this work, prepaid subscribers’ behavior are defined with their RFM data and some additional features, such as usage, call center and refill transactions. We model the churn behavior using Pareto/NBD model and with two benchmark models: a logistic regression model based on RFM data, and a logistic regression model based on the additional features. Pareto/NBD model is a crucial step in calculating customer lifetime value (CLV) and aliveness of the customers. If Pareto/NBD model proves to be a valid approach, then a mobile operator can define valuable prepaid subscribers using this and decide on the actions for these customers, such as suggesting customized offers. | |
| 19 | @inproceedings{Macis2015ATI,  title={A TV-based ICT Platform for Active Ageing, Tele-care and Social Networking},  author={Silvia Macis and Daniela Loi and Danilo Pani and Wil Rijnen and Luigi Raffo},  booktitle={ICT4AgeingWell},  year={2015},  url={https://api.semanticscholar. org/CorpusID:15697364}  } | The modern society is dealing with a progressive increase of the elderly population. The development of services for social inclusion and independent living is of paramount importance to enable the elderly to live in their homes autonomously as long as possible. Such a solution paves the way to a sustainable social and economic model where older adults develop self-confidence and promote their participation to the community life. This paper presents the hardware/software framework of a novel ICT system for active ageing support, which combines the potentialities of broadband internet services to the simplicity of TV use. User research in three European countries allowed to define several important services (healthcare, home monitoring, shopping, communication and social inclusion) to be provided through the developed platform. Its modularity, supported by the App paradigm, enables easy customization and future developments. | |
| 20 | @article{Zhou2016SimulationOA,  title={Simulation of a Tele-operated Task under Human-Robot Shared Control},  author={Longjiang Zhou and Keng Peng Tee and Zhiyong Huang},  journal={Proceedings of the Fourth International Conference on Human Agent Interaction},  year={2016},  url={https://api.semanticscholar. org/CorpusID:18185887}  } | This poster presents simulation of a tele-operated shared controlled robot task that is integrated with a generic simulator of RADOE (Robot Application Development and Operating Environment). A customized and extendable Rviz interface plugin is designed and applied to import models, do simulation, enable real robot operation, and communicate with other projects by clicking related buttons. In the simulation process, the robot model in the simulator is controlled and visualized by human operator using an Omega 7 haptic device and automatic method, i.e., the shared control. After the simulation is conducted and satisfied, the system will send back a signal to the real robot system to execute the operation task; otherwise, simulation of the shared control process will be continued until satisfaction. We provide a simulation of a drawing task on the surface of a sphere. | |
| 21 | @inproceedings{Fernndez2014TheRO,  title={The Relevance of Providing Useful and Personalized Information to Therapists and Caregivers in Tele},  author={Juan Manuel Fern{\'a}ndez and Marc Sol{\`a} and Alexander Steblin and Eloisa Vargiu and Felip Miralles},  booktitle={DART@AI\*IA},  year={2014},  url={https://api.semanticscholar. org/CorpusID:13802502}  } | Nowadays, filtering and analyzing data coming from Tele\* (i.e. telemedicine, telerehabiliation, telemonitoring, telecare, and teleassistance) systems is becoming more and more relevant. In fact, those systems gather a lot of data coming from patients through wearable, domotic, and environmental sensors, as well as questionnaires and interviews. The role of therapists and care givers is essential for remotely assisting the corresponding patients. Thus, intelligent solutions able to understand all those data and process them to keep therapists and caregivers aware about their assisted persons are needed. Moreover, friendly and useful tools for accessing and visualizing those data must be provided to therapists and caregivers. In this chapter, we present a generic Tele\* solution that, in principle, may be customized to whatever kind of real scenarios to give a continuous and efficient support to therapists and caregivers. The aim of the proposed solution is to be as flexible as possible in order to be able to provide telerehabilitation, telemonitoring, teleassistance or a conjunction of them, depending on the real situation. Three customizations of the generic platform are also presented. | |
| 22 | @inproceedings{Brady2016ALC,  title={A Low Cost Desktop Robot and Tele-Presence Device for Interactive Speech Research},  author={Michael Connolly Brady},  booktitle={Interspeech},  year={2016},  url={https://api.semanticscholar .org/CorpusID:30049130}  } | In building robotic systems that interact with people through speech, many robotics engineers are obliged to treat artificial speech recognition and synthesis as a black-box problem best left to speech engineers to solve. Yet speech engineers today typically do not have access to the kinds of expensive robots needed for this development. Progress on the human-robot speech interface thus suffers from something of a diffusion of responsibility. In an attempt to remedy the situation, we have developed a low-cost interactive embodied speech device. The device is constructed from off-the-shelf components and from 3D-printed and laser-cut parts. We make the files for the 3D and laser-cut parts freely available for download. In addition to offering basic assembled devices and kits for self-assembly, we provide an assembly guide and a shopping list of components a user will need in order to build, maintain, and customize their own device. We supply a basic software framework (in both Matlab and in C/C++), and template code for a ROS node for interfacing with the device. The idea is to establish a standard and accessible hardware platform with an open-source foundation for the sharing of ideas and research. | |
| 23 | @article{Kushawaha2022ASO,  title={A Study of Artificial Neural Network and Its Implementation from Scratch},  author={Nilay Kushawaha and Ankhi Roy},  journal={Journal of Robotics and Automation Research},  year={2022},  url={https://api.semanticscholar .org/CorpusID:251691257}  } | One of the major problems in the field of artificial intelligence (AI) is the use of machine learning model as a black box, even though it might be helpful in a few cases but understanding the internal structure and the operating mechanism will assist the user to tweak the variables in a more efficient and productive manner. In this paper we have introduced the working of an artificial neural network (ANN) by taking the example of a three layered neural network. The entire mathematics behind the working of neural network along with the different evaluation metrics required to assess the performance of the model are discussed in this paper. We have also created a custom neural network from scratch and compared it with the keras based model on three different datasets - susy dataset [8], cardiovascular dataset [6], churn dataset [7]. The results obtained demonstrates that the overall performance of both the models are almost identical which gives an idea that it is possible to train a neural network from scratch without the use of any framework. | |
| 24 | @article{Seo2023ImprovingSM,  title={Improving Shopping Mall Revenue by Real-Time Customized Digital Coupon Issuance},  author={Daeho Seo and Yongmin Yoo},  journal={IEEE Access},  year={2023},  volume={11},  pages={7924-7932},  url={https://api.semanticscholar. org/CorpusID:256227902}  } | With the development of big data and deep learning technology, big data and deep learning technology have also been applied to the marketing field, which was a part of business administration. Customer churn management is one of the most important areas of marketing. In this paper, we proposed a method to prevent customer churn and increase purchase conversion rate by issuing customized discount coupons to customers with high churn rate based on big data in real time. After segmenting customer segments with two-dimensional segment analysis, a real-time churn rate estimation model based on clickstream data was generated for each segment. After that, we issued customized coupons to our customers. Finally, we tested the conversion rate and sales growth. A two-dimensional cluster analysis-based churn rate estimation combined with a recommendation system was found to be significantly more useful than the respective simple models. Using this proposed model, it is possible to increase sales by automatically estimating the customer’s churn probability and shopping propensity without the burden of marketing costs in the online shopping mall. | |
| 25 | @article{Mendes2023OpenDPMHAF,  title={OpenDPMH: A Framework for Developing Mobile Sensing Applications of Digital Phenotyping},  author={Jean P M Mendes and Francisco Silva and Andr{\'e} Cardoso and Ivan Rodrigues de Moura and Luciano Reis Coutinho and Davi Viana and Markus Endler and Ariel Soares Teles},  journal={2023 IEEE 36th International Symposium on Computer-Based Medical Systems (CBMS)},  year={2023},  pages={198-203},  url={https://api.semanticscholar. org/CorpusID:259958875}  } | Digital Phenotyping of Mental Health (DPMH) aims to passively collect data from ubiquitous devices to be used as evidence in the process of diagnosis, treatment, and monitoring. Literature presents different sensing mobile applications for digital phenotyping, however they are not extensible and can not be customized for use in other research. In this paper, we propose OpenDPMH, a framework for developing mobile sensing applications able to collect contextual data in order to produce useful user information that represent situations of interest for mental health professionals and researchers, such as human behaviors and habits. Our solution is extensible and reusable, as it allows the inclusion of modules for collecting and processing new raw context data with features for data distribution. By implementing a case study, we demonstrate that OpenDPMH is suitable for the development of DPMH mobile applications. Moreover, we carried out experiments to evaluate the energy consumption on smartphones, which demonstrate a low battery cost to run applications developed using the proposed framework. | |
| 26 | @article{Roslan2021iNutritionAppMA,  title={iNutritionApp: Mobile Application for Nutrition Monitoring using FatSecret API},  author={Nurul Nabilah Roslan and Muhammad Nabil Fikri Jamaluddin and Alif Faisal Ibrahim and Shukor Sanim Mohd Fauzi and T. R. Razak and Ray Adderley Jm Gining},  journal={Journal of Computing Research and Innovation},  year={2021},  url={https://api.semanticscholar .org/CorpusID:239666819}  } | Mobile application is a software designed to run on smartphones, tablet computers, and other mobile devices. The growing millions of users who are using mobile applications have contributed to an increase in the development of the mobile applications for enterprises, education, the social network and healthcare. Widely used healthcare application nowadays includes general health and wellness, tele-medicine, personal coaching and consultation, medical record tracking, custom reminders and various health management apps. This paper presents a mobile application related to general health and wellness named as iNutritionApp for providing nutrient information, tracking nutrition and calorie intakes with the integration of FatSecret API. The API provides nutritional information based on type of food provided by user. The application was intended to overcome manual calorie intake calculations and assists diet plan. Development methodology of this mobile application utilizes three phases that are system requirements, system design and development, as well as testing. Technology acceptance model with three parts were conducted with 30 respondents by evaluating the developed mobile application through questionnaires. Results of the testing showed that perceived of usefulness (PU) part achieved highest mean score compared to perceived ease of use (PEOU) that include user interface design and navigation parts. Therefore, features and functionality offered by the iNutritionApp is proven to be useful for user in tracking calorie intakes and provide access to nutritional information. | |
| 27 | @inproceedings{Cotter2019OnMS,  title={On Making Stochastic Classifiers Deterministic},  author={Andrew Cotter and Maya R. Gupta and Harikrishna Narasimhan},  booktitle={Neural Information Processing Systems},  year={2019},  url={https://api.semanticscholar. org/CorpusID:202770212}  } | Stochastic classifiers arise in a number of machine learning problems, and have become especially prominent of late, as they often result from constrained optimization problems, e.g. for fairness, churn, or custom losses. Despite their utility, the inherent randomness of stochastic classifiers may cause them to be problematic to use in practice for a variety of practical reasons. In this paper, we attempt to answer the theoretical question of how well a stochastic classifier can be approximated by a deterministic one, and compare several different approaches, proving lower and upper bounds. We also experimentally investigate the pros and cons of these methods, not only in regard to how successfully each deterministic classifier approximates the original stochastic classifier, but also in terms of how well each addresses the other issues that can make stochastic classifiers undesirable. | |
| 28 | @article{Aujeszky2020AFF,  title={A Framework for Thermographic Material Characterization Using Multichannel Neural Network},  author={Tam{\'a}s Aujeszky and Georgios Korres and Mohamad A. Eid},  journal={IEEE Transactions on Instrumentation and Measurement},  year={2020},  volume={69},  pages={7061-7071},  url={https://api.semanticscholar. org/CorpusID:216220938}  } | Research on material characterization has received an increasing amount of attention recently. In several application scenarios, it is essential to effectively estimate the physical properties of objects without coming into contact with them (e.g., tele-operated or autonomous robotics). This article presents the Haptic Eye: a framework using active thermography that uses a custom multichannel neural network approach to perform classification between samples and regression toward their thermal properties. This neural network structure is uniquely suited for effective processing of thermographic data. The framework is realized, implemented, and evaluated with a set of ten samples with diverse thermal/physical properties. Experimental results on a realization of the framework validate this approach, with a 92.20% classification accuracy using multichannel neural network with majority vote, as well as more than 99.6% R 2 -fit with respect to three different thermal properties, namely thermal conductivity, thermal diffusivity, and thermal effusivity (which is very useful to define thermal exchange during physical interaction). | |
| 29 | @article{Aujeszky2019EstimatingWO,  title={Estimating Weight of Unknown Objects Using Active Thermography},  author={Tam{\'a}s Aujeszky and Georgios Korres and Mohamad A. Eid and Farshad Khorrami},  journal={Robotics},  year={2019},  volume={8},  pages={92},  url={https://api.semanticscholar .org/CorpusID:208011745}  } | Successful manipulation of unknown objects requires an understanding of their physical properties. Infrared thermography has the potential to provide real-time, contactless material characterization for unknown objects. In this paper, we propose an approach that utilizes active thermography and custom multi-channel neural networks to perform classification between samples and regression towards the density property. With the help of an off-the-shelf technology to estimate the volume of the object, the proposed approach is capable of estimating the weight of the unknown object. We show the efficacy of the infrared thermography approach to a set of ten commonly used materials to achieve a 99.1% R2-fit for predicted versus actual density values. The system can be used with tele-operated or autonomous robots to optimize grasping techniques for unknown objects without touching them. | |
| 30 | @article{Behrendt2021OrchestrationAS,  title={Orchestration and Situation Awareness in an Assistance System for Assembly Tasks},  author={Wernher Behrendt and Felix Strohmeier},  journal={IFIP Advances in Information and Communication Technology},  year={2021},  url={https://api.semanticscholar .org/CorpusID:234201407}  } | We report on the design, specification and implementation of a situation awareness module used for assistive systems in manufacturing, in the context of Industry 4.0. A recent survey of research done in Germany and Europe, concerning assistive technology in industry shows a very high potential for “intelligent assistance” by combining smart sensors, networking and AI. While the state of the art concerning actual technology in industrial use points more towards userfriendly, speech-based interaction with personal assistants for information retrieval (typically of in-house documentation), the research presented here addresses an enterprise-level assistance system that is supported by a number of specialized Assistance Units that can be customized to the end users’ specifications and that range from tutoring systems to tele-robotics. Key to the approach is situation awareness, which is achieved through a combination of a-priori, task knowledge modelling and dynamic situation assessment on the basis of observation streams coming from sensors, cameras and microphones. The paper describes a working fragment of the industrial task description language and its extensions to cover also the triggering of assistive interventions when the observation modules have sent data that warrants such interventions. | |
| 31 | @inproceedings{Anonymous2021NavigatingHT,  title={Navigating Healthcare Through Challenging Times - Proceedings of dHealth 2021 - Health Informatics Meets Digital Health, Virtual Conference, May 11-12, 2021},  author={Anonymous},  booktitle={dHealth},  year={2021},  url={https://api.semanticscholar .org/CorpusID:239980260}  } | As one digital intervention for informal caregivers, the ‘Angehörigenampel’ (caregivers’ traffic-light) was developed, which is able to assess the physical and psychological burden of caregivers. This can help to counteract the health effects of caregiving burden early on before it is too late. The development of the digital intervention as a WordPress-plugin was kept generic so that it can easily be adapted to other languages on further websites. The ‘intervention as a plugin’ approach demonstrates an easy and flexible way of deploying eHealth interventions to other service providers, especially from other countries. The implementation barriers for other service providers are low enough for them to be able to easily integrate the eHealth intervention on their website, enabling more caregivers to benefit from the disseminated eHealth intervention. | |
|  |  | |
|  |  | |
|  |  | |
|  |  | |
|  |  | |
|  |  | |
|  |  | |
|  |  | |
|  |  | |
|  |  | |
|  |  | |
|  |  | |
|  |  | |
|  |  |  | |

# **I) Giới thiệu bài toán(Introduce)**

In the telecommunications industry, data from phone calls, messages, internet data and other services is becoming an important source of information, providing extensive insights into user behavior and performance. of the network. For telecommunications service providers, analyzing and extracting insights from this data not only helps them better understand customer needs but also helps them optimize business strategies and improve performance. service quality.( Trong ngành viễn thông, dữ liệu từ các cuộc gọi điện thoại, tin nhắn, dữ liệu internet và các dịch vụ khác đang trở thành một nguồn thông tin quan trọng, cung cấp thông tin sâu rộng về hành vi của người dùng và hiệu suất của mạng lưới. Đối với các nhà cung cấp dịch vụ viễn thông, việc phân tích và trích xuất insights từ dữ liệu này không chỉ giúp họ hiểu rõ hơn về nhu cầu của khách hàng mà còn giúp họ tối ưu hóa các chiến lược kinh doanh và cải thiện chất lượng dịch vụ.)

Problem description:

Input data: Includes information about phone calls such as calling phone number, receiving phone number, call time, location, and other technical information such as signal level, call type, and information about packages.

Target:

Analyze customer behavior: Predict calling trends, analyze calling and receiving habits, identify popular and uncommon calling patterns.

Measure network performance: Monitor call failure rates, average connection times, and geographic call distribution to detect and resolve technical issues.

Detect fraud and spam: Look for predictive models and detect fraudulent or spam calls.

Meaning and application:

Better understand customer behavior and needs: Analyzing call data can help telecommunications service providers better understand customers' calling habits, thereby creating new services and solutions. suitable package.

Improve service quality: By monitoring and evaluating network performance from call data, providers can detect and resolve technical issues quickly, enhancing the user experience. use.

Cybersecurity protection: Call data analysis can help detect and prevent scam calls, spam, and activities that threaten the security of users' personal information.

Project scope:

Collect, preprocess, and explore call data from telecommunications systems.

Apply big data processing and machine learning techniques to analyze and extract insights from data.

Build fraud detection and prediction models from call data.

Create reports and graphs to present analysis results and evaluate model performance.

Analyzing data from telecommunications systems not only helps us better understand customer behavior but also plays an important role in improving service and network management. We hope that applying Apache Spark in this problem will bring valuable insights and make a positive contribution to the telecommunications industry.

(Mô tả bài toán:

Dữ liệu đầu vào: Bao gồm thông tin về các cuộc gọi điện thoại như số điện thoại gọi, số điện thoại nhận, thời gian gọi, địa điểm và các thông tin kỹ thuật khác như mức độ tín hiệu, loại cuộc gọi, và thông tin về gói cước.

Mục tiêu:

Phân tích hành vi của khách hàng: Dự đoán xu hướng gọi điện thoại, phân tích thói quen gọi và nhận cuộc gọi, xác định các mô hình gọi điện thoại phổ biến và không phổ biến.

Đo lường hiệu suất mạng lưới: Theo dõi tỷ lệ gọi thất bại, thời gian kết nối trung bình, và sự phân phối của cuộc gọi theo vùng địa lý để phát hiện và giải quyết vấn đề kỹ thuật.

Phát hiện gian lận và spam: Tìm kiếm các mô hình dự đoán và phát hiện các cuộc gọi lừa đảo hoặc gọi spam.

Ý nghĩa và ứng dụng:

Hiểu rõ hơn về hành vi và nhu cầu của khách hàng: Phân tích dữ liệu cuộc gọi có thể giúp nhà cung cấp dịch vụ viễn thông hiểu rõ hơn về thói quen gọi điện thoại của khách hàng, từ đó tạo ra các dịch vụ và gói cước phù hợp.

Cải thiện chất lượng dịch vụ: Bằng cách theo dõi và đánh giá hiệu suất mạng lưới từ dữ liệu cuộc gọi, các nhà cung cấp có thể phát hiện và giải quyết các vấn đề kỹ thuật một cách nhanh chóng, tăng cường trải nghiệm người dùng.

Bảo vệ an ninh mạng: Phân tích dữ liệu cuộc gọi có thể giúp phát hiện và ngăn chặn các cuộc gọi lừa đảo, spam và các hoạt động đe dọa bảo mật thông tin cá nhân của người dùng.

Phạm vi dự án:

Thu thập, tiền xử lý và khám phá dữ liệu cuộc gọi từ hệ thống viễn thông.

Áp dụng các kỹ thuật xử lý dữ liệu lớn và học máy để phân tích và trích xuất insights từ dữ liệu.

Xây dựng các mô hình dự đoán và phát hiện gian lận từ dữ liệu cuộc gọi.

Tạo ra các báo cáo và đồ thị minh họa để trình bày kết quả phân tích và đánh giá hiệu suất mô hình.

Bài toán phân tích dữ liệu từ hệ thống viễn thông không chỉ giúp chúng ta hiểu rõ hơn về hành vi của khách hàng mà còn đóng vai trò quan trọng trong việc cải thiện dịch vụ và quản lý mạng lưới. Chúng ta hy vọng rằng việc áp dụng Apache Spark trong bài toán này sẽ mang lại những insights giá trị và đóng góp tích cực vào ngành viễn thông.)

**II) Nghiên cứu liên quan(Related research)**

In the telecommunications field, many important studies have been conducted with the aim of researching and developing data analysis methods from telecommunications data sources such as phone calls, text messages and internet data. . Here are some related studies that we can refer to: ( Trong lĩnh vực viễn thông, có nhiều nghiên cứu quan trọng đã được thực hiện với mục đích nghiên cứu và phát triển các phương pháp phân tích dữ liệu từ các nguồn dữ liệu viễn thông như cuộc gọi điện thoại, tin nhắn và dữ liệu internet. Dưới đây là một số nghiên cứu liên quan mà chúng ta có thể tham khảo: )

"Big data analytics for user mobility and call activities prediction in telecom: A survey" (2020) by Hou et al.: This study provides an overview of big data analytics methods and techniques for mobility prediction user transfer and calling activity in the telecommunications industry. The author analyzed machine learning methods and predictive models to understand and predict customer behavior. ("Big data analytics for user mobility and call activities prediction in telecom: A survey" (2020) của Hou et al.: Nghiên cứu này cung cấp một tổng quan về các phương pháp và kỹ thuật phân tích dữ liệu lớn để dự đoán di chuyển người dùng và hoạt động gọi trong ngành viễn thông. Tác giả đã phân tích các phương pháp học máy và mô hình dự đoán để hiểu và dự đoán hành vi của khách hàng.)[1]

"Telecommunications data for the humanitarian community: A systematic review" (2021) by Avsigol et al.: This study focuses on exploring the potential of telecommunications data in supporting humanitarian activities. The author evaluated the applications of telematics data in predicting and responding to emergencies and emergency situations, such as post-disaster relief response. ("Telecommunications data for the humanitarian community: A systematic review" (2021) của Avsigol et al.: Nghiên cứu này tập trung vào việc khám phá tiềm năng của dữ liệu viễn thông trong việc hỗ trợ các hoạt động nhân đạo. Tác giả đã đánh giá các ứng dụng của dữ liệu viễn thông trong việc dự đoán và phản ứng đối với các tình huống khẩn cấp và tình trạng khẩn cấp, như phản ứng cứu trợ sau thiên tai.)[2]

"Predictive analytics for customer churn in the telecommunications industry: A systematic literature review" (2021) by Zeng et al.: This study focuses on applying predictive methods to prevent and minimize customer churn products in the telecommunications industry. The author has synthesized customer churn prediction methods and models from previous studies to provide an overview of this field. ("Predictive analytics for customer churn in the telecommunications industry: A systematic literature review" (2021) của Zeng et al.: Nghiên cứu này tập trung vào việc áp dụng các phương pháp dự đoán để phòng ngừa và giảm thiểu sự rời bỏ của khách hàng trong ngành viễn thông. Tác giả đã tổng hợp các phương pháp và mô hình dự đoán khách hàng rời bỏ từ các nghiên cứu trước đây để đưa ra cái nhìn tổng quan về lĩnh vực này.)[3]

The above studies provide diverse perspectives on the application of telecommunications data and data analysis methods in this field. We can take advantage of the knowledge and experience from these studies to apply to our problems and create effective solutions. (Những nghiên cứu trên cung cấp các góc nhìn đa dạng về ứng dụng của dữ liệu viễn thông và các phương pháp phân tích dữ liệu trong lĩnh vực này. Chúng ta có thể tận dụng những kiến thức và kinh nghiệm từ những nghiên cứu này để áp dụng vào bài toán của chúng ta và tạo ra những giải pháp hiệu quả.)