

# LITERATURE SURVEY

**AUTHOR NAME:** Kristijan Cincar

**PROPOSED TOPIC:** Hospital Management and Scheduling with Multi Agents Approach

**DATE & YEAR :** October 29-30,2020

**PROBLEM:**

This paper proposes a patient scheduling system based on Multi-agent systems to balance workload of medical staff while minimizing patient waiting time. A model that predicts length of stay and patient satisfaction could lead to shorter hospital stays, reduced risk of infection, improved quality of treatment, and increased hospital earnings. Automation of hospital management systems can lead to more efficient and satisfactory programs. The Hospital Model is a system model that prioritizes patients based on priority, consists of agents and manages interactions between them. Hospitals are complex organizations with multiple stakeholders, and the system must account for these differences and adapt to dynamic demand.

Hospital management and scheduling with a multi-agent approach is a challenge due to the limited availability of resources. This paper proposes a solution to balance the workload of medical staff while minimizing patient waiting time for nonemergency appointments. patients must be involved in the design and implementation process to ensure successful adoption of a new system. Resistance to change is another problem in hospital management and scheduling with a multi-agent approach. It is essential to consider stakeholders' needs and preferences, involve them in the design and implementation process, and provide training and support to ensure successful adoption of the new system.

**AUTHOR NAME :** Ritu Chauhan, Eiad Yafi

**TOPIC :** Big Data Analytics for Prediction Modelling in Healthcare Databases

**DATE & YEAR :** MAY 28,2021

**PROBLEM:**

This paper provides an overview of big data analytics applications in healthcare and government. It discusses the challenges of big data analytics and how it can be used to improve healthcare outcomes. Data mining techniques such as Decision Trees, Nearest Neighbor methods, Rule Induction, Genetic Algorithms, Applied Artificial Neural Networks and Regression are used in cancer research to detect hidden knowledge from big databases. The incidence of cancer in young adults rose to 1,516 cases, increasing to 35,261 in the age group of 50-54, and dropping to 13,24 in 85+ years. Example: Breast

cancer is the leading cause of cancer death among women in Australia, with the most affected being middle-aged females. Early diagnosis and treatment can help reduce incidences and mortality.

Big data analytics is a powerful tool for prediction modeling in healthcare databases. It can extract insights from large and complex datasets, such as EHRs and medical imaging data, to predict disease outcomes and improve patient care. However, it presents significant challenges such as data privacy and security concerns, ethical considerations, and technical barriers. However, with the promise of improving healthcare outcomes and reducing costs, big data analytics is a rapidly growing field that has the potential to revolutionize healthcare delivery and management.

**AUTHOR NAME :** Rahul Katarya, Sajal Jain.

**TOPIC :** Exploration of Big Data Analytics in Healthcare Analytics

**DATE & YEAR :** June 16,2021

**PROBLEM :**

Big data is a large and complex data set that requires analyzing and cost-effective management for extraction. It has five v's: variety, volume, value, velocity, and veracity. The data in healthcare is in terabytes (10<sup>12</sup> bytes), petabytes (10<sup>15</sup> bytes), and zettabyte (10<sup>21</sup> bytes) and still growing exponentially. The health sector faces many challenges, such as the collection and processing of structured and unstructured data, the selection of the right tool, and the slow incorporation of new technologies. Cost also plays a role in success. Data in the healthcare sector can be structured, semi-structured or unstructured, with structured data having a defined model and semi-structured data having organizational properties. Clinical data is unstructured and includes external data such as biometrics, insurance policies, and financial data.

Hadoop is an open-source software framework that solves the storage problem of massive amounts of unstructured data in healthcare. It consists of four stages: data collection, cleaning, transformation into a research database, and predictions obtained using big data tools and algorithms. This paper will help policymakers, hospitals, patients, and pharmaceuticals companies choose the best tools and algorithms for performing the analysis.

Big data analytics in healthcare analytics has the potential to revolutionize healthcare, but it presents challenges such as quality and interoperability of data, privacy and security regulations, and analytics models. However, it has the potential to identify high-risk patients and predict disease outcomes.

**AUTHOR NAME :** EUNBI KIM , KAP SU HAN , TAESU CHEONG , SUNG WOO LEE , JOONYUPEUN , AND SU JIN KIM

**PROPOSED TOPIC :** Analysis on Benefits and Costs of Machine Learning-Based Early Hospitalization Prediction

**DATE & YEAR :** March 21,2022

**PROBLEM:**

Emergency department (ED) overcrowding is a severe problem in the health sector worldwide due to a discrepancy between medical demands and resource supplies. Big data analytics has the potential to transform healthcare by improving patient outcomes and optimizing operations. This study aims to identify a model that accurately predicts ED patients who are hospitalized to inpatient beds at an early stage of ED stay and estimate the quantitative effects of hospitalization predictions on EDs and wards. Predictive models can help identify patients who are more likely to be hospitalized using readily available information, such as vital signs.

This text discusses three studies that used data from 52,037 children to predict hospitalization. Horng et al. used data from 230,936 ED patients to predict their diseases. Ram et al. used Twitter and Google data to predict the number of daily ED visits of asthmatics. This study investigates its use in hospitalization predictions for ED patients. It makes four major contributions: it shows that high-accuracy hospitalization predictions are possible using data on less than 50,000 patients, it analyzes hospitalization predictions for all ED patients without confining the group under study to patient characteristics, and it uses Natural gradient boost (NGBoost) to derive a good prediction performance from a small set of data. It also shows quantitative effects of hospitalization predictions on EDs and wards. The most important idea is to select a model with a high true-positive rate to reduce ED overcrowding. Accuracy and AUC are important factors to consider.

**AUTHOR NAME:** Ganga Naga Saroj Bandi, Dr Tamarana.Srinivasa Rao, Sayyad Saadiq Ali

**PROPOSED TOPIC:** Cybersecurity and Data Privacy in the Cloudlet for Preliminary Healthcare Big Data Analytics

**DATE & YEAR :** September 10,2020

**PROBLEM:**

A cyber physical system (CPS) is the backbone of 4IR, integrating automation and knowledge for manufacturing. Interoperability, virtualization, decentralization, real-time capability, service orientation, and modularity are design principles of 4IR in Healthcare, with cyber security and data privacy challenges. Cloudlets are fog computing virtual components that form a micro datacenter nearby mobile devices to facilitate offloading data and computation. Cloud computing is used for medical big data analytics tasks, such as A/B testing, association rule learning, and machine learning. Cyber-physical devices are defined as devices that interact with the physical world through sensing and actuation, with security, privacy, safety, reliability, and resilience being high-level concerns.

Distributed cloudlets attached to imaging sensors can be used to conduct preliminary analytics on visual big data, useful for behavior modification suggestions and remote robotic surgery. Healthcare organizations need to be aware of the potential risks and take appropriate measures to protect sensitive data and ensure compliance with regulations when using cloudlets for preliminary healthcare big data analytics. To mitigate these risks, healthcare organizations need to implement strong cybersecurity measures and carefully select cloudlet providers with a strong track record of security and compliance.

**AUTHOR NAME:** Ahmed I. Taloba , Rasha M. Abd El-Aziz , Huda M. Alshanbari, and Abdal-Aziz H. El-Bagoury

**TOPIC NAME :** Estimation and Prediction of Hospitalization and Medical Care Costs Using Regression in Machine Learning

**DATE & YEAR :** March 2,2022

**PROBLEM:**

The incidence of overweight and obesity has increased significantly in recent decades, leading to an increased incidence of chronic diseases. Modelling uses machine-learning methods and regression is a popular predictive analytic model used to forecast healthcare costs for BMI. Weka 3.8.1 was used to forecast Tw-DRG49702 patients using random forest, support vector machines, Naive Bayesian, C4.5 decision tree, and logistic regression approaches. A deep learning based natural language processing (NLP) method was tested to forecast cost-reflecting weights and per-episode DRGs on 2

cohorts. Linear regression is a machine learning statistical analysis technique used to find linear correlations between two or more responses and predictive variables.

The most commonly used measurements to estimate the performance of a linear regression are the root mean square error (RMSE), mean absolute error (MAE), and mean square error (MSE). Data exploration and preparation is the next step in machine learning, and 80% of efforts are dedicated to data. Tsuyama Chuo Hospital provided healthcare record information to forecast a patient's healthcare costs for the coming year. The IEVREG framework is an accessible framework that allows for the calculation of the proportion (mass) of every element of information in the testing phase L. The proposed linear regression model seeks out individuals with identical genders, BMIs, children, and smoking statuses, while ignoring age.

Average annual rates and costs of consultations, tests, and prescription items were estimated by BMI category at the time of recruitment. Subgroups defined by alcohol consumption, socioeconomic status, smoking level, educational qualifications, and strenuous exercise were estimated using the chi square test. The method is effective in estimating healthcare costs with an accuracy rate of 97.89%.