LITERATURE SURVEY

1. CHRONIC KIDNEY DISEASE (CKD) PREDICTION USING DATA MINING TECHNIQUES

Abstract: In the past decade, the rapid growth of digital data and global accessibility through the modern internet has seen a massive rise in machine learning research. In proportion to it, the medical data has also seen a massive surge of expansion. With the availability of structured clinical data, researchers have attracted scores to study clinical disease detection automation with machine learning and data mining. Chronic Kidney Disease (CKD), also known as the renal disorder, has been such a field of study for quite some time now. Therefore, our research aims to study the automated detection of chronic kidney disease using several machine learning classifiers with clinical data. The purpose of this research work is to diagnose kidney disease using a number of machine learning algorithms such as the Support Vector Machine (SVM) and the Bayesian Network (BN) and to select the most effective one to assess the extent of CKD patients. The amount of expertise in the medical field in relation to CKD is limited. Many patients have to wait a long to get their test results. The experience of medical staff is declining in value. Upon retirement, new employees replace them. It helps professional doctors or medical staff in their diagnosis of CKD. This paper's primary purpose is to present a clear view of Chronic Kidney Disease (CKD), its symptoms, and the process of early detection that may help humanity be safe from this life-threatening disease.

2. **Real-Time Smart Attendance System using Face Recognition Techniques.**

Abstract: The management of the attendance can be a great burden on the teachers if it is done by
hand. To resolve this problem, smart and auto attendance management system is being utilized. But authentication is an important issue in this system. The smart attendance system
is generally executed with the help of biometrics. Face recognition is one of the biometric methods to improve this system. Being a prime feature of biometric verification, facial recognition is being used enormously in several such applications, like video monitoring ...

3. Supervised Machine Learning Algorithms for Credit Card Fraud Detection:

Abstract : Fraud is one of the most major ethical issues in Credit card industry. The main Purpose of our paper is to identify the Credit card fraud and provide a reasonable Solution to the fraud. Frauds caused by Credit Cards have costs consumers and banks billions of dollars globally. Even after numerous mechanisms to stop fraud, fraudsters are continuously trying to find new ways and tricks to commit fraud. Fraud detection is of immense importance in banking field and finance related companies. We are going to apply artificial neural network for detection purposes. Thus in order to stop it we will provide a solution which will not only detect fraud but will detect it before it happening. Our system will learn from past committed fraud in order to detect new frauds. Mining algorithms had been applied to detect fraud but did not performed well. In our paper we are implementing machine learning algorithms to detect fraud in credit card transactions. The paper utilizes

the supervised learning algorithms which are implemented on a dataset from kaggle which was highly skewed and imbalance. We balanced the set by robust scalar to have a 51 percent non fraud cases and 49 percent fraud cases. Logistic regression, random forest, decision tree and KNN has been implemented and further learning curves are displayed which shows which algorithm has the best ability to perform.

The metrics used for evaluation are accuracy, specificity, precision and sensitivity and a comparative chart is established which displays the comparative analysis of these supervised learning algorithms.

4. Utilizing a Data analytics project to increase student awareness of contemporary global issues.

   Abstract: Computational thinking is irrefutably a musthave skill in today's digital world. This is why it is a requisite course in most college curricula. Computational thinking is a broad term encompassing fundamental concepts of computing such as formulating problems and expressing their solutions in computational steps that can be processed by a computer. The ability to organize and analyze data is also a part of computational thinking. Following this paradigm, educational institutions have adopted various methods to teach computational thinking including programming. This paper will explore a specific project within the computational thinking course taught as an introductory course in all major programs at the Higher Colleges of Technology in the United Arab Emirates. The intention of the paper is to demonstrate how a project-based data analytics assessment in such a course can be used to foster in students, a greater awareness of a contemporary and critical global issue such as waste management. To examine the effectiveness of this project, the paper documents and evaluates the work of a group of students who carried out data analysis on secondary data collected from a notable website on "waste management". The methodology implemented for this task is developed based on a prescribed course syllabus and assessment structure which promotes student-centered and inter-disciplinary learning.

5. Prediction and Diagnosis of Heart Disease Patients using Data Mining Technique .

Abstract: We are living in a post modern era and there are tremendous changes happening to our daily routines which make an impact on our health positively and negatively. As a result of
these changes various kind of diseases are enormously increased. Especially, heart disease has become more common these days. The life of people is at a risk. Variation in
Blood pressure, sugar, pulse rate etc. can lead to cardiovascular diseases that include narrowed or blocked blood vessels. It may causes Heart failure, Aneurysm, Peripheral artery ...

6. Fake news detection system using data science .

Abstract: Initially, the platform must be constructed in accordance with the data format associated with false and authentic news. The implemented programmes must be synchronised with the data structure during the design phase. The bogus database displays no news channel names, but the genuine dataset displays individual headquarters for each station. Manipulating the concept of dataset fraudulent channels are exploiting an unregistered news portal. As a result, using the original dataset, one may compare and explicitly identify them. In this venture, we are using LS-TM Recurrent Neural Network using (Long Short Term Memory) to forecast fake news because there is a large amount of fake news in all types of media such as social media or news media, and the author is training LS-TM' Genuine' and 'Fake' news data were used to train a neural network. We found FAKE NEWS messages on Twitter on the internet.