## **Artificial Intelligence and Machine Learning**

Project Report

Semester-IV (Batch-2022)

**Case Study**: - **Support Vector Machine (SVM) for Cell Classification**

[**https://drive.google.com/file/d/1vIO6D5uS8iL71jECyr0wcg3Ffx07G4PC/view?usp=sharing**](https://drive.google.com/file/d/1vIO6D5uS8iL71jECyr0wcg3Ffx07G4PC/view?usp=sharing)

A red and white sign

Description automatically generated with low confidence

**Supervised By: Submitted By:**

Rajeev Thakur Yatin Kalra

Roll Number: -2210990983

Group - 14

**Department of Computer Science and Engineering**

**Chitkara University Institute of Engineering & Technology,**

**Chitkara University, Punjab**

**Case Study: Support Vector Machine (SVM) for Cell Classification**

**Objective**:

The objective of this analysis is to utilize Support Vector Machine (SVM) to classify human cell records as benign or malignant based on certain features.

**Methods Used:**

* Pandas library
* NumPy library
* Matplotlib library
* Scikit-learn library

**Methods:**

* read\_csv(): Reads a CSV file and converts it into a data frame.
* train\_test\_split(): Splits the dataset into train and test sets.
* SVC(): Initializes the SVM classifier.
* fit(): Trains the SVM classifier.
* predict(): Predicts the labels using the trained SVM classifier.
* classification\_report(): Generates a classification report including precision, recall, and F1-score.

**Analysis Steps:**

* *Data Loading and Exploration:*
  + Read the cell samples dataset using read\_csv() method.
  + Display the tail of the dataframe to ensure completeness.
  + Check the shape, size, and count of the dataframe.
  + Visualize the distribution of classes (benign and malignant).
* Data Preprocessing:
  + Select relevant features and remove unwanted columns.
* Model Building and Evaluation:
  + Split the data into train and test sets using train\_test\_split() method.
  + Initialize the SVM classifier with a linear kernel.
  + Train the SVM classifier using the training data.
  + Predict labels for the test data.
  + Evaluate the performance of the classifier using classification\_report() method.

**Report:**

* Data Loading and Exploration:
  + Displayed the tail of the dataframe to ensure completeness.
  + Checked the shape, size, and count of the dataframe.
  + Visualized the distribution of classes to understand the data distribution.
* Data Preprocessing:
  + Selected relevant features and removed unwanted columns to prepare the data for modeling.
* Model Building and Evaluation:
  + Utilized Support Vector Machine (SVM) with a linear kernel to build a classification model.
  + Split the data into train and test sets to evaluate the model's performance.
  + Trained the SVM classifier using the training data.
  + Predicted labels for the test data using the trained model.
  + Generated a classification report including precision, recall, and F1-score to assess the model's performance.

This structured approach provides a clear overview of the steps involved in building and evaluating an SVM model for cell classification, facilitating easy replication and understanding of the analysis.