Human Activity Recognition using Smartphone Dataset

Fundamentals of Machine Learning (INT-254)

END TERM REPORT

Submitted by

Abhishek Kumar

Registration no.: 12014108

Section: KM117

Roll Number: B48

Submitted to

Dr Pawan Kumar Mall

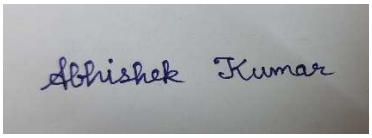
School of Computer Science and Engineering

Lovely Professional University, Jalandhar

November, 2022

Student Declaration

This is to declare that this report has been written by me. No part of the report is copied from other sources. All information included from other sources have been duly acknowledged. I aver that if any part of the report is found to be copied, I shall take full responsibility for it.



Abhishek Kumar

Reg. no.: 12014108

Roll no.: RKM117B48

Lovely Professional University, Phagwara, Punjab

14th November, 2022

Acknowledgement

It is my pleasure to acknowledge the contributions of all who have helped and supported me during the making of this project.

I’m thankful to my faculty Dr Pawan Kumar Mall who taught me from the very beginning of the day about basic of fundamental of Machine Learning up to the level that I’m able to create this project successfully. It would not have been possible without him.

I’m thankful to my friends and classmates who create a good environment of learning from each other.

I’m also thankful to our parents, elders and all family members for their blessing, motivation and inspiration throughout my work. They have always provided me a high moral support and contributed in all possible ways in completion of this Capstone report.

Abstract

The goal of this project is to build a Machine Learning model and a signal processing pipeline in offline mode capable of processing signals collected using smart phone inertial sensors and producing useful datasets will be used as inputs of a machine learning model capable of recognizing some of human daily activities (sitting, walking, exercising, etc.) included in the dataset with a low error rate.

The signal processing pipeline and the final model could be used as a good source of information about user's daily activities needed by remote monitoring systems for various purposes.

Table of Contents

Introduction

Machine learning is almost everywhere nowadays. It's become more and more necessary day by day. From the recommendation of what to buy to recognizing a person, robotics, machine learning is everywhere. So, in this project, I'll demonstrate a sample ML model that can be used to identify user activity based on data from the user's smartphone sensors.

Since 1870 a large growth in human life expectancy has been observed in Europe. This growth has expanded in the whole world principally due to the great achievements in health care field. As a result, the proportion of elderly people is rapidly increasing. Aging people in general lives in isolated conditions. In addition to that some of them are not capable to live normally and take advantages from health care facilities services. Building remote monitoring systems for elderly patients who live alone or without permanent caretaking will improve their quality of life. For better decision making these remote monitoring systems needs some regular and trustful information about patients.

What is Machine Learning?

Machine learning is about learning to predict something or extracting knowledge from data. ML is a part of artificial intelligence. ML algorithms build a model based on sample data or known as training data and based upon the training data the algorithm can predict something on new data.

Categories of Machine Learning

Machine Learning can be broadly classified into two distinct methodologies: Supervised and Unsupervised Machine Learning. We'll take a brief look at both in this section.

Supervised Machine Learning

Supervised machine learning are types of machine learning that are trained on well-labelled training data. Labelled data means the training data is already tagged with the correct output.

Unsupervised Machine Learning

Unlike supervised learning, unsupervised learning doesn’t have any tagged data. It learned patterns from untagged data. Basically, it creates a group of objects based on the input data/features.

Applications of Machine Learning

1. **Speech Recognition**: Speech recognition uses NLP (Natural Language Processing) to process human speech into written format and vice versa. Some examples are – Google Assistant, Alexa, Siri.
2. **Recommendation Engine**: Using the past behaviour of a human's search data the recommendation engine can produce new data to cross-sell products to customers. For example – Amazon product recommendations, Spotify music recommendations.
3. **Chatbots**: Chatbots are used to give customer services without any human agent. It takes questions from users and based on the question it gives an answer as a response.

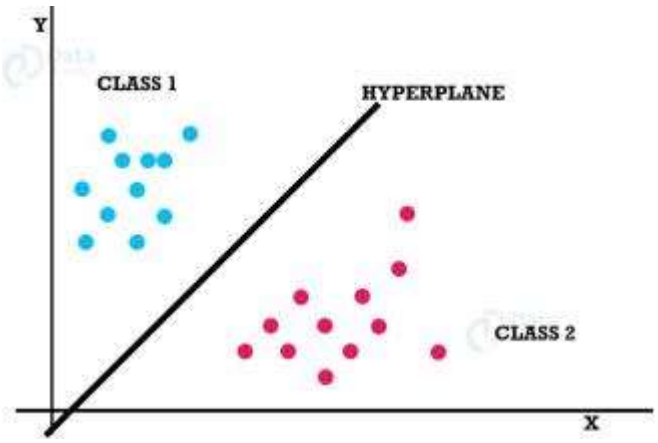
Human Activity Detection using Smartphone Dataset

In this report, I'll demonstrate the use of "Support Vector Machine" to implement a Machine Learning Model capable of detecting user activity from data of their smartphone sensors, such as accelerometer or gyroscope.

Support Vector Machine

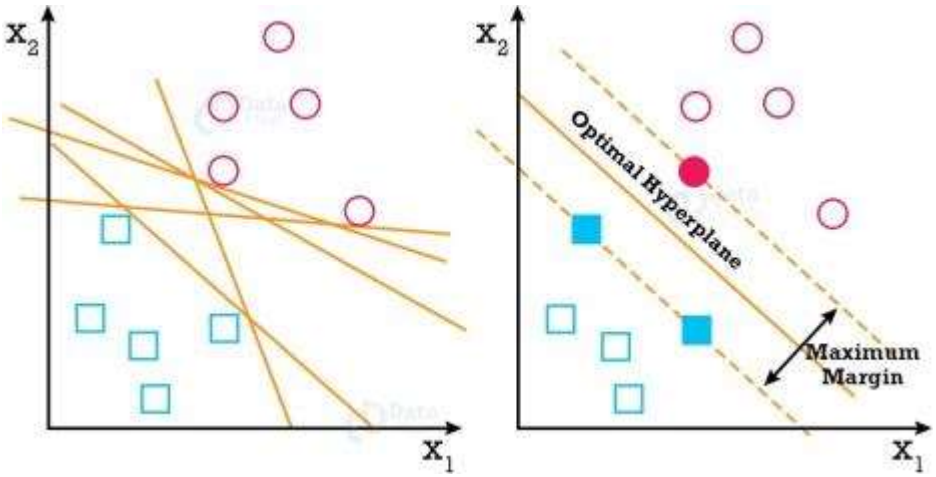
A support vector machine (also known as a support vector network) is a supervised machine learning algorithm that analyses data for classification and regression. SVMs are one of the most robust classifications methods.

SVM approximates a separate line (Hyperplane) between two classes.



SVM algorithm finds the points closest to the line from both classes. These points are known as support vectors. Then it computes the distance between the line and support vectors.

This distance is called the margin. The main goal is to maximize the margin. The hyperplane which has the maximum margin is known as the optimal hyperplane.



SVM mainly supports binary classification natively. For multiclass classification, it separates the data for binary classification and utilizes the same principle by breaking down multi-classification problems into multiple binary classification problems.

Software Requirements

This project requires **Python 3.5** and the following Python libraries installed:

1. Python 3.5
2. NumPy
3. SciPy
4. Pandas
5. Matplotlib
6. StatsModel
7. Spectrum

The code and dataset for the model can be downloaded from my GitHub repository: https://github.com/abhi-kr-2100/Human-Activity-Recognition-Using-Smartphone-Dataset.

Steps to detect human activity from smartphone sensors:

1. Conduct an experiment to collect data. Sample data sets are already available on the web. The dataset used for this project is based on an experiment described below:
    a. The experiments were carried out with a group of 30 volunteers within an age bracket of 19-48 years. They performed a protocol of activities composed of six basic activities: three static postures (standing, sitting, lying) and three dynamic activities (walking, walking downstairs and walking upstairs).
    b. The experiment also included postural transitions that occurred between the static postures. These are: stand-to-sit, sit-to-stand, sit-to-lie, lie-to-sit, stand-to-lie, and lie-to-stand.
    c. All the participants were wearing a smartphone (Samsung Galaxy S II) on the waist during the experiment execution. They captured 3-axial linear acceleration and 3-axial angular velocity at a constant rate of 50Hz using the embedded accelerometer and gyroscope of the device.

    d. The experiments were video-recorded to label the data manually
2. Feature Selection:
    a. Subsequently, the body linear acceleration and angular velocity were derived in time to obtain Jerk signals. Also, the magnitude of these three-dimensional signals was calculated using the Euclidean norm.
    b. Finally, a Fast Fourier Transform (FFT) was applied to some of these signals.
3. Windowing: These Signals were then sampled in fixed-width sliding windows of 2.56 sec and 50% overlap.
4. Features Generation: From each sampled window, a vector of 561 features was obtained by calculating variables from the time and frequency domain.
5. Training: The Support Vector Machine model was obtained after training using the SVM algorithm on the 561-feature vectors.
6. Running: Finally, the file “Machine Learning Part.ipynb” can be run inside a Jupyter notebook to get the results of this experiment.

Summary

In this project, I learned to train our own supervised machine learning model using Support Vector Machines. Through this project based on determining human activity from smartphone sensor data, we learned about machine learning, data analysis, data visualization, model creation, etc.