**TITLE**

**ABSTRACT**

**ACKNOWLEDGEMENTS**

**TABLE OF CONTENTS**

**LIST OF TABLES AND FIGURES**

**CHAPTER 1: Introduction**

**Section 1:1 Introduction to the problem**

On the basis of different GPX files uploaded by different users round the globe, we are trying to estimate or build up a model so that it can empower us to predict sports activities in different geo-tracks. Lot of IOT devices like smart phones, smart watches are being regularly used to capture data while performing these kind of activities. Fast life and high expectancy of life made us serious about getting involved in fitness programs. This project is an attempt to not only predict activities for a certain geo track but also plotting an association between similar tracks. According to specific fitness goal, one can choose similar tracks, available as per his choice of city or country. As we tried to cluster tracks as per the difficulty level, the project will be helpful to choose other tracks to increase the fitness level as per the selected sport activity.

[[implement different clustering methods and find efficiency.

Predictive analysis of different sports activities. (academic paper)]]

**Section 1:2 Project Objectives**

two or three specific objectives

1. Grouping GPX files as per their characteristics and prediction of activity performed to generate the file.
2. Building a web platform to enable users to choose tracks as per their fitness program. For a certain activity, the tracks will be grouped as per different level of competency levels. Users will be able to choose lower or higher or same level of tracks. The selection can be made as per their neighbourhood or city or any country basis.
3. The project may help various government bodies to set their priorities to sanction or complete projects in building tracks suitable for sports like running, hiking, cycling etc. If they can estimate the utilization of a prospective route and predict the kind of activities will be performed on them, will certainly help the organizations to plan for cost estimation and feature build up.

**Section 1:3 Overall Solution**

Not detailed specification of methodology

Prior to come up with a solution, it is essential to list down what kind of information we are accumulating from the GPX files. Each of the GPX files consist of a series of nodes and each of the node holds four types of data. Timestamp, altitude, longitude & latitude. The combination of latitude and longitude pin points the location whereas the altitude states about the elevation details. The timestamp field provides information related to date and time. The starting node usually signifies the starting point of any journey. Whereas the last point usually lets us know about the end point of the journey. Having said that, we will consider the start and end points based upon time fields available within the file. Only these four fields are not good enough to find out the required clustering among different routes based upon the performed activity. We will make an attempt to derive some more fields from these original fields. The main objective to find out these derived fields is that it will provide us better understandability about the difficulty of a certain track. For an example, if there is a sharp inclination in the way then it is likely that a hiker or runner will face challenges and it will result in decrease of his/her speed. To figure out the change in elevation between successive nodes, we will plot the difference in ‘DeltaElev’ field. One interesting point is that the gadget being used, does not capture on regular intervals. And the speed of the person varies at different point of time as well. So, we would like to calculate the time differences between each two nodes in ‘TimeDiff’ field. From the combination of latitude and longitude fields, we get to realize the ‘Point’ field. From the ‘point’ field, we can compute the distance crossed between each successive nodes and the result is inserted in ‘GeoPointsDist’ field. One of the most significant thing is to know how straight a path is. More bend found in a track, more likely it causes decrease in speed and less distance is covered as well. For activities like cycling, motor-cycling are very much impacted by the change of this field, ‘Angle’. For the start and end node we considered the angle as zero. To calculate the angle of any given point, we collected it’s prior and subsequent node for reference. If it is evident that all the three points are in a straight line, then also we consider its angle as zero.

**[For the time being: Data Manipulation is written here.]**

Each of the result set is grouped with the respective filename and summary statistics is formed according to that.

When we are done with

**Section 1.4 Outline of the thesis**

**CHAPTER 2: Background and Related Work**

**Section 2:1 Related Literature and Other Works**

**Youtube videos: gist**

**Section 2:2 Outline of contributions**

**CHAPTER 3: Design and Solution Overview**

To be completed today

**CHAPTER 4: Analysis and Evaluation**

**CHAPTER 5: Conclusions and Future Work.**

**Section 5:1 Summary of Thesis:**

**Section 5:2 Overall Evaluation**

**Section 5:3 Future Work**

**REFERENCES/BIBLIOGRAPHY**

**APPENDICES**

**Summarising academic papers**