Real-World Visual Data Analysis of Garmin FIT Files for Training Development

University of Lincoln, School of Computer Science
Nathan Jones – 19696498@students.lincoln.ac.uk

Abstract—This research proposal poster informs the audience of the motivation behind the research, along with the main aims and objectives of the research which details how they will be carried out within the research. These are crucial metrics that will be evaluated at the end of the research when determining the success of the research's findings, the research methods for the research have been critically discussed with a Gantt Chart visualising the timeline of the research, that show key milestones and goals of the research.

Introduction

Over the past two decades the cycling world has revolutionised the way in which data is recorded during recreational, training and race activities from the introduction of cycling computers such as the Garmin Edge series paired with devices such as heart rate monitors, power meters, cadence sensors, speed sensors, temperature sensors and GPS connectivity produce crucial data that is waiting to be utilised. This wealth of data can be analysed to understand training patterns and improve performance of individuals through training development based upon data analytics.

One area of data analytics that has taken off in the past 15 years is the use of visual data analysis or data visualisation, this is the graphical representation of data through graphs, charts and, heat maps to produce visual elements that allow for patterns and trends within the data to be identified faster. Software solutions such as Strava, VeloViewer and TrainingPeaks have taken this visual data analysis approach to the market for athletes to analyse their data through performance metrics [1] such as suffer score, weighted average power, intensity, training load and power curve. However, over the past few years these once free and fully accessible solutions have become subscription services with many of these features locked behind paywalls, making it impossible for individuals to analyse their data the same way as before.

Academic research within the domain of visual data analytics around cycling data and Garmin FIT files for training development is limited with the bulk of the research conducted surrounding machine learning [2] and professional male athletes [3][4], instead of the wider population. This gap in analytical research is part of the driving force behind the proposed research along with the change in attitude within the market towards the analytical software solutions becoming subscription services.

The Real-World Visual Data Analysis of Garmin FIT Files for Training Development research sets out to produce comprehensive academic literature along with a software solution that will allow athletes of all levels to analyse their data. This data analysis solution will utilise visual data analytics to produce easy to understand visuals with training trends, performance analysis including correlations, performance metrics and other relevant metrics within cycling. This software solution will allow users to upload their own recreational, training or race FIT files into the software and their data will be analysed along with any trends within their past and present data. This analysis can then be utilised in developing training plans based upon these trends and patterns within the data to improve their performance without the need for a subscription.

Aims and Objectives

The research aims to fulfil a set number of aims and objectives to be deemed successful in its research, these aims create the basis of the research and are the foundations of the objectives. The the main aims of the research are detailed below and are then expanded within the objective section.

Aims

- Read and Decode Garmin's Proprietary FIT File Format
- Compute Linear Correlational Analysis with Data Visualisation Outputs
- Produce Performance Analysis and Metrics using Data Visualisation Tools
- Produce Training and Performance Trends Based Upon Past and Present Data
- Produce a Software Solution Containing All of the Above for User Interaction
- Produce Clear and Easy to Interpret Visual Analytical Data Plots
- Produce Academic Literature Documenting and Evaluating the Research

Objectives

- Data Analysis Objectives:
  - Identify Correlations Between Metrics (Heart Rate, Power, Speed, Cadence, etc.)
  - Utilisation of Linear Correlation Techniques such as Correlation Coefficient Values
  - Analyse and Identify Patterns within the Data Using Data Visualisation Techniques
  - Map Trends within the Data and Possible Future Trajectories
  - Identify Key Areas within the Data to Improve Performance Metrics
- Software Solution Objectives:
  - Read and Decode Garmin's Proprietary FIT Files into Usable Dictionaries and Arrays
  - Individual User Profiles Through a Login System and Database
  - Access to Multiple Garmin FIT Files Unique to Each User
  - Easy to Navigate with a Clean and Streamlined User Interface
  - Data Visualisation in Clean Understandable Plots to Identify Trends and/or Patterns in the Data
- Literature Objectives:
  - Comprehensive and Critical Report that Details the Research from Start to Finish
  - Comprehensive and Critical Literature Review that Explores the Current Academic Literature and Any Previous Research with this Domain of Data Analytics
  - Detailed Development Logs Throughout the Process
  - Critical and Detailed Evaluations on the Whole Research Throughout
  - Detailed Findings and Conclusions with Recommendations for Future Research

The purpose of the research is to investigate the correlations between performance metrics along with utilising data analytics to identify trends and patterns within the data that can then be used to develop new training plans that improve the athlete's overall performance. This research will extensively discuss and evaluate all methods, software developments, findings and conclusions in a detailed manner so that if any further research within the domain is required, other researchers have a clear detailed basis to launch from. Along with the academic literature a software solution will be developed in conjunction with the above aims and objectives, this solution will be a free and widely available product that will allow users to create their own accounts along with importing their own Garmin FIT files for data analysis to allow for training development through trends and patterns within their data.

Research Methods

The research will use a correlation analysis method throughout, this will allow for the strength and direction between variables within the research to be assessed throughout the research by providing valuable insights into patterns, trends and potential connections within the data. This is key in this research that is looking at training development from past data and the potential of predicting future performance metrics based off these correlations. For this research Garmin FIT files will be obtained from a Garmin Edge 510 paired with a Garmin Heart Rate Monitor, Garmin Cadence Sensor and Garmin Vector 3 Power Meter Pedals. This will provide the necessary performance data required to provide analysis for training development. Python will be the language utilised to visualise the data collected from the Garmin device along with the correlations, trends and patterns within the data to improve the performance metrics through training developments, external python modules will be utilised for the data analysis, visualisation and GUI such as SASPY, Matplotlib, and Tkinter.

Research Plan

The research plan has been formed in a Gantt Chart to visualise key milestones along with the development and research goals within the given timeframe of the research that has been predefined (June 2023-September 2023).
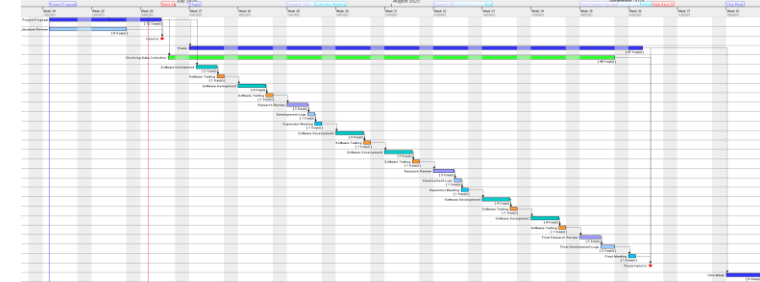
References

[1] K. Wamsley, "Optimal power-based cycling pacing strategies for Strava segments - ProQuest," *www.proquest.com*. May, 2014 https://www.proquest.com/docview/1527593761?pqorigsite=gscholar&fromopenview=true

[2] T. Jin, "Predicting Pace Based on Previous Training Runs" *CS229: Machine Learning, Stanford University*, Dec. 12, 2014. [Online]. Available: http://cs229.stanford.edu/proj2014/Tiffany%20Jin,%20Predicting%20Pace%20Based%20on%20Previous%20Training%20Runs.pdf

[3] T. van Erp, D. Sanders, and J. J. de Koning, "Training Characteristics of Male and Female Professional Road Cyclists: A 4-Year Retrospective Analysis" *International Journal of Sports Physiology and Performance*, pp. 1–7, 2019, doi: https://doi.org/10.1123/ijspp.2019-0320

[4] J. Pinot and F. Grappe, "The Record Power Profile to Assess Performance in Elite Cyclists," *International Journal of Sports Medicine*, vol. 32, no. 11, pp. 839–844, Nov. 2011, doi: https://doi.org/10.1055/s-0031-1279773

*A complete list of references is available upon request