CSU44061 Short Project Proposal: Weather and Cycling in Dublin

Conall Tuohy 18320949, Tanmay Kaushik 18308341, Liam Bartsch 18330219

November 16, 2021

Contents

1   Motivation

What problem are you tackling?

Cycling in Dublin seems to be a great option no matter the weather because it's fairly easy to get from one side of the city to the other with the numerous cycling lanes in the city (Cycle Dublin in a day), but it seems well known that the number of cyclists will decrease after the harsher conditions that rain introduces.

We are tackling the problem of whether or not, and to what degree, weather affects cycling in Dublin from 2019 to 2021, to test these claims and find out more information on what else might influence cyclists. We plan to find out what factors cyclists consciously or unconsciously use to gauge whether they should cycle by looking at their trends over a large time. Similar research has been done but it was focused on a smaller time frame, focusing on different things and the depth of the research did not satisfy our curiosity and interest in the domain. We plan to introduce a number of features of the weather to see what aspects of the weather most change the volume of cyclists in Dublin. From this, we might also discover some interesting trends between the number of cyclist and the weather, that could be used in further areas of research.

2   Dataset

What data will you use and how will you collect it?

We decided to combine the three datasets from Cycle Count from January 2019 to October 2021. We also chose to use the Cycle Count's locations. We were unable to find data from the Phoenix Park Weather station for our selected time frame (this may have been due to the fact that the station was closed and that it wasn't quality checked) which is why we decided to use data from the closest reliable station which is Dublin Airport Weather station Dublin Airport Weather Station. We will have to filter the raw data that is within the CSV files that were available to download from these websites and parse it to our needs.

3   Method

What machine learning techniques are you planning to apply or improve upon?

We are planning to use Linear Regression, Lasso regression, Ridge regression and *knn* models to try to tie a relationship between the input features of the weather and the output feature the cyclist count.

4 Intended experiments

What experiments are you planning to run? How do you plan to evaluate your machine learning algorithm?

We intend to train a linear regression model, a Lasso regression model, a Ridge regression model as well as a *knn* model, to see the weights these models give to the different features for our post-analysis. Once we have trained our models we will also use some baseline models and compare which models perform the best, given some pre-selected testing data. Once we have evaluated the performance of each model and compared them to each other we intend to analyse the data and the models themselves to see if we can uncover other relationships between the weather and cycling in Dublin.