School of Computing  
CA326 Year 3 Project Proposal Form

**SECTION A**

Project Title \_\_\_FormulaML\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Student 1 Name \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ ID Number \_\_\_\_\_\_\_\_\_\_\_

Kealan O’Connor

18348311

Student 2 Name \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ ID Number \_\_\_\_\_\_\_\_\_\_\_

18424692

Peter Browne

Student 3 Name \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ ID Number \_\_\_\_\_\_\_\_\_\_\_

*(A third team member is exceptional and requires detailed justification.)*

Staff Member Consulted \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Dr. Michael Scriney

Project Description (1-2 pages):

**Description:**

Our Project, FormulaML, is a Machine Learning based web application that will analyse and break down large collections of data and statistics from previous Formula 1 races in order to predict the possible outcomes of future races and to allow our users to understand and think like a Formula 1 team. Formula 1 is the highest level of motorsport and is by its very nature one of the most technical sports there is. Races are won and lost not by seconds but by hundredths and thousandths of a second and so every minute detail can make a huge difference to the outcome of a race. In recent years the importance of complex simulations and data analytics have played a huge factor in the strategies that teams choose to implement and the drivers they choose to race with. Under normal circumstances a normal fan of the sport could not hope to keep up with everything the teams are factoring in however our aim for this project is to give even the most casual fans a tool that will allow them to think, analyse and predict a race like a formula 1 team.

Our project will gather data using an Opensource API called ERGAST which allows users to access the FIA archives dating all the way back to the 1950s and create a mySQL database from all of this data. We will then Analyse this data using machine learning techniques such as Linear Regression and Monte Carlo Simulations in order to devise a system of ratings and scenarios that will form the basis of our predictions. Once our predictions for each driver, team and engine manufacturer are made we plan to display this data in a number of ways on our website in a way that will give users valuable insight while also being user-friendly and easy to understand.

**Division of work:**

We both have worked on projects together before and so establishing a dynamic between us is simple. We plan to split the work evenly between us at each stage of the project using an agile approach that will allow us to divide tasks according to the particular stage or focus of the project. We also plan to use Pair-Programming as we have experience with this system and it is a great way to ensure work is evenly spread and it allows us to both play into our strengths and weakness’ or gaps in technical knowledge.

**Programming Languages:**

* Python 3 – This will be our primary language for our backend.
* JavaScript – This will primarily be used for our frontend and website.
* MySQL – This will be used for our database and data gathering using the ERGAST API.

**Programming Tools:**

* Vscode – This will be our primary programming environment as we both regularly use it and it allows us to pair programme remotely.
* Django – This will be used to manage our backend.
* NodeJS – This will be used for our frontend and server.
* ERGAST API – This is an open source API we will use to help gather our data.

**Learning Challenges:**While both of us have had an interest in machine learning and data analysis this will be our first large project implementing it, however we look forward to learning from it. We are also unfamiliar with using the Django framework but it will be a useful skill to add to our portfolio. We are learning about REST servers and NodeJS in some of our other modules so while not completely new to us this project will give us a good opportunity to put what we’ve learned to the test.

**Hardware/software platform**:

We will be developing this project on a Linux environment on a PC as this the environment we are most comfortable working in.

**Special Requirements:**

For the purpose of this project we do not require any specific or special requirements.