1. **Motivation & Background**

The area of injury prediction within the NFL is a crucial piece of study that can have significant implications on players, teams and the league as a whole. Annually, there is an estimated 1.2million injuries across all levels of American football[2], this shows the necessity for an all-encompassing approach to injury prediction in the NFL. These injuries not only have effects on the players themselves, resulting in long-term health problems or even possibly retirement, but it can also have significant impacts on team performance and their subsequent profitability. The NFL generates roughly $16bn annually[11], indicating the sheer amount of money that is in the NFL, and having their best players out through injury can have a significant knock-on effect on these revenues. If the best players are always out injured performances fall, leading to a lack of interest and ultimately a hit to both match day revenue and commercial revenue, if fans both stop showing up at matches or watching on television.

More importantly, injury prevention is a matter of player safety and career longevity. Given the heavy contact nature of the NFL, big hits and their subsequent injuries are common place and so it’s important to try and limit these possible career threatening moments. The question of head injuries and concussions has long since been a topic of debate across various sports the world over, none more so than the NFL and the heavy hitting that is so prevalent in today’s game can be cause for some more long-term and possibly significant impacts on players health. The idea of CTE and its effects on NFL players is a somewhat more recent area of study however, despite that around 96% of deceased ex-NFL players have been diagnosed with the disease[7]. This shows an urgent issue that needs to be solved to ensure that former players can continue to live long and happy lives post-retirement.

Unfortunately, the task of predicting injuries is a complex one, with various possible factors it means it is necessary to conduct large scale research to even just discover which variables would have any sort of influence on injuries. The factor selection is also incredibly vast from weather type to play style all of these can have varying levels of importance an any one facet of injury risk making the research all the more complex. Also, given the highly specialised nature of these machine learning models the results become incredibly difficult to describe to people without the necessary domain knowledge, this calls for the use of more explainable AI that will allow for easier interpretability of how the models work and their results for people without this aforementioned knowledge. Data or lack thereof is another major issue for creating these sorts of models, the area of injury prediction is relatively novel when considering the use of machine learning to achieve it. Because of this there is not a consensus database, that comprehensively catalogues all the necessary information to conduct this sort of research. Data imbalance is also a problem for developing a machine learning model, the instances of injuries are hugely underrepresented in any NFL play-by-play database, meaning that it’s essential that sampling techniques take place to make sure the model is able to predict injuries in any meaningful way for analysis.

This all being said, it’s incredibly important that research towards large scale injury prediction within the NFL is conducted so as to limit long-term life-altering injuries for its players. This research would allow all stakeholders from coaches to medical staff and even the players themselves to make well-informed decisions about their health and recovery plans. Personalised recovery and training would be able to become a regular part of preseason sessions as well as throughout the season improving player safety and extending careers. With better injury prediction models, the NFL would not only gain a competitive and financial advantage but would also become world leaders in the area of injury prediction and prevention.