

# 36-460/660 Final Project Report

Alex Cheng, Melody Wang, Liz Chu, Kevin Ren

April 23rd, 2024

## Contents

<b>Introduction</b>	<b>1</b>
<b>Data</b>	<b>1</b>
<b>Methods</b>	<b>2</b>
<b>Results</b>	<b>2</b>
<b>Discussion</b>	<b>2</b>

## Introduction

*Describes the problem you're tackling, why it is interesting or important. Gives a brief summary of your results (though most of your results will be at the end of your report).*

The problem of playcalling is among the greatest challenges in sports coaching, particularly in the sport of American football. Particularly poor calls – e.g. the Seahawk's decision to pass rather than run with 1 yard to go in the closing seconds of Super Bowl XLIX – represent snap decisions that can cause franchises and players alike to lose critical games and ultimately millions of dollars. As such, in order to avoid such outcomes coaches must be very smart about what kinds of plays to run in different scenarios, depending on the previous plays that they have called along with the context of the game. In this project we hope to answer the age-old question of what plays a football coach should call, given the types of the previous three plays in the drive, along with contextual information about the field position such as yards-to-go until a touchdown, the strength of the offensive team in terms of their passing and rushing abilities, and the difference in score in the game. We ultimately find that our fitted models on the given covariates perform strongly on both in-distribution and out-of-distribution data, compared to a naive baseline model.

## Data

*Describes the data you're using in detail, where you accessed it, along with relevant exploratory data analysis (EDA). You should also include descriptions of any major pre-processing steps.*

**TODO:** we should probably include the code that got us `runrunrun.csv` somewhere in our code submission

## Methods

*Describes the modeling techniques you chose, their assumptions, justifications for why they are appropriate for the problem, and your plan for comparison/evaluation approaches. Additionally, you will need to describe how you will quantify uncertainty for your estimates of interest, with sufficient descriptions of the approach and justification for why it's appropriate for your data and problem of interest.*

## Results

*Describes your results. This can include tables and plots showing your results, as well as text describing how your models worked and the appropriate interpretations of the relevant output. I do not want you to write out the textbook interpretations of all model coefficients! I only want you to interpret the output that is relevant for your question of interest that is framed in the introduction.*

## Discussion

*Gives your conclusions and summarizes what you have learned with regards to your question of interest. Are there any limitations with the approaches you used? What do you think are the next steps to follow-up your project?*