Final Project -Notes

Tuesday, January 25, 2022 7:31 PM

 # Week of 1/30/22

 # FINAL Project: Predicting College Football Outcome (Win/Lose) Based on First -Half Statistics

OBJECTIVE: Taking conference college football games from 2008 -2013, can we predict the winner of the each game based on first -half statistics?

SUMMARY: Blue team chose the topic because it satisfied the interest of the team, datasets were readily available via Kaggle, and it p rovides all the key elements desired by the exercise, the ability to utilize machine learning to attempt to predict a binary outcome win/lose based on qualities (rush, pass, penalties et al) of a large variety of games. In addition there is an expansive capability to bring in additional data that may contribute to outcome and subsequent visual representation and interaction to a user interested in predicting game outcomes at the half.

Source data will come from Kaggle and includes, but may not be limited to college football: game stats, conference details, e t al. Questions that may be answered

1) Can a game be predicted win/lose based on first-half stats, and to what level of accuracy?
2) How close does predicted meet actual results?
3) What values contribute most heavily to the predicted outcome?
4) What other values may contribute to win/lose?
5) Is the model used in machine learning the best model for this exercise?

TEAM, TASKS, OWNERSHIP

Name	Tasks	Notes	Role
Ambrea Curtis	GitHub Visual Tool Presentation	• GITHUB: Owner of team repository, owns pulling branch data into main (after team confers) • VISUAL TOOL: Evaluating a couple visual output opportunities, prototyping how to present to user. First priority is to tell the output story and enabling user to have the required interaction. • Presentation: Key stakeholder in evaluating, performing and presenting select portions of project	Square-responsible for repository
Brian Moazen	Python Machine Learning Presentation	• Python: Bring in core Kaggle data to determine first 1/2 stats. CSV output will then require SQL joins and extraction for 'clean lake' csv. • Machine Learning: (Model TBD) -ultimate goal is to predict win/lose based on 1/2 stats and run against actual bahavior. • Presentation: Key stakeholder in evaluating, performing and presenting select portions of project	Triangle: Responsible for machine learning
Ben Peyton	Visual Output Presentation	• VISUAL TOOL: Evaluating a couple visual output opportunities, prototyping how to present to user • Presentation: Key stakeholder in evaluating, performing and presenting select portions of project	X: Responsible for Visual and for confirming technologies used
Lauren Lodl	PgAdmin/Posgress SQL Presentation Notes	• DATABASE/SQL: Take 'murky-lake' Python data into PGADMIN/Posgress SQL, perform join(s) and produce 'clean-lake' csv for machine learning and visual output(s) • Presentation: Key stakeholder in evaluating, performing and presenting select portions of project	Circle: Responsible for database

Primary Communication Methods:

1) **Tuesday/Thursday are our status meetings where we will answer**
    a. What have we done since we last met
    b. What are we going to accomplish by the time we meet again
    c. Are there any blockers
        i. If there are blockers we will spend time as a team to discover what we can do to get past them
2) We will communicate via Slack for times in between
3) Team is available to meet at different times as needed on zoom if we are behind

Conflict Management:

1) Owner of the task listed above will ultimately decide how to proceed.
    a. As a team we will try to negotiate a solution that works for all parties if there is a project conflict.
2) If there is interpersonal conflict it is recommended that the persons with issues take some time offline to try to resolve sp ecific items in question

Future Notes-Placeholder:

Week of 2/6
What we are working on between 2/1 to 2/6

Team Member	**What working on since we met last**	**What working on between now and next meeting**	**Blockers**
Ambrea	Merged files in GitHub, preliminary map visuals	Delete out of GitHub (play 2005 and 1st half stats function), draft README from notes, brainstorming details for visual (line/bar graphs). Create prototypes off of Kaggle data	no
Brian	Performed prototyped modeling shared with group on 2/1	Utilize gold copy csv and try additional models for greater predictability/sensitivity. Meet on Saturday hours for expanded capabilities	no
Ben	Visual brainstorming	Working on the popups on the map, HTML groundwork. Meet on Saturday hours for expanded capabilities. Mock up of user input UI.	no
Lauren	Created Posgress (local version) database to produce joins, ETL import and export complete	remove gold copy data set with half score >14 ; add integer field to identify hometeam (1/0) and surface Updated notes out on the site. Meet on Saturday for expanded capability. Move databases (Python/VSCode) into SQLite, database for prediction input model and database for future java inputs.	no

Week of 2/13

Team Member	**What working on since we met last**	**What working on between now and next meeting**	**Blockers**
			Blockers **Additional time with TA will help identify how to integrate UI/backend/ and visuals

			appropriately
Ambrea	Fine tune visual support	Continue to draft visual support	no
Brian	Integrate machine learning into sql database (in/out)	Finalize python function for machine learning Draft plots for presentation in PNG.	no
Ben	Adjust user input UI to align with final user inputs.	Get the css file complete,prototype machine learning integration into web apps.	no
Lauren	Fine tune database capability in SqLlite to support Java entries/machine learning	Draft Start first slides Objective, questions, High-level of the backend; SQL (ERD connected), ETL, database	no

Week of 2/20