SPANISH LEAGUE : BET RECOMMENDER

WHAT ?

IDEA

- CREATE A RECOMMENDENDER TO OPTIMISE THE ODDS TO WIN A BET AND/OR IDENTIFY OPPORTUNITIES.

- THE RECOMMENDATOR WILL RETRIEVE THE PERCENTAGE OF SUCCESS FOR DETERMINATED BETS ,HOW THE QUOTAS ARE OVER OR UNDERPAYED TO FIND 'BARGAINS' AND WHAT IS THE BEST OPTION(S) FOR EACH ROUND OF THE FOOTBALL SPANISH LEAGUE.

- BESIDES I WILL MAKE THESE RECOMMENDATIONS USING 3 RISK DIFFERENT STRATEGIES: CONSERVATIVE, 'HUNTER OF BARGAINS' OR MIXED.

WHY ?

THE HOUSE ALWAYS WINS

- I ALWAYS WONDERED IF THERE IS ' A BEST STRATEGY' TO OPTIMISE MY BETS AND HOW MUCH I CAN REDUCE THE LUCK/UNCERTAINTY WHEN I GAMBLE.

- BLACK BOX : HOW THE QUOTAS ARE CALCULATED? WHAT ARE THE ODDS ASSOCIATED TO THE EVENTS RELATED TO THOSE QUOTAS ? ARE THEY PAYING A FAIR PRICE? CAN WE TAKE ADVANTAGE OF SOME 'OVERESTIMATED' QUOTAS?

- ON THE WAY AND EVEN IN THE CASE THE RECOMMENDATOR IS NOT AS SUCCESFUL AS I EXPECT I WILL HAVE EXPLORED WHICH STRATEGIES ARE THE BEST ONES AND WILL HAVE UNDERSTOOD IN THAT CASE WHY 'THE HOUSE ALWAYS WINS'.

- ALL THE WEBSITES RECOMMENDING BETS BASED ON PERCENTAGES DON'T SHOW HOW THESE % ARE CALCULATED*, WHAT'S THE BEST OPTION(S) FOR THAT ROUND , DON'T SHOW THE POTENTIAL 'BARGAINS' AND ARE NOT SPLIT INTO DIFFERENT 'RISKY PROFILES' (STILL EXPLORING THE WORLD OUT THERE)

HOW ?

SCOPE & NEEDED DATA

- TO SIMPLIFY AND AS A FIRST APPROACH I WILL ONLY ANALYSE SOME SCENARIOS (1, X, 2, BOTHS TEAMS WILL SCORE) FOR THE SPANISH LEAGUE.

- I WILL NEED TO COLLECT AS MUCH INFORMATION AS I CAN FROM THE GAMES FOR THE LAST SEASONES*. ALSO I NEED TO COMPLETE IT WITH MATCHES FROM OTHER COMPETITIONS SINCE COULD AFFECT TO THE RESULTS IN THE LEAGUE.

- FINALLY I WILL NEED TO RETRIEVE THE PAST** AND THE UPCOMING*** QUOTAS.

* HYSTORICAL MATCHES RESULTS IN THE SPANISH LEAGUE

** BET EXPLORER

*** BET365 API

METHODOLOGY ? (I)

DATASETS

- FIRST I NEED TO COMPLETE A VERY WELL QUALIFIED DATABASE WITH AS MANY VARIABLES AS POSIBLE SINCE THAT WILL BE THE FOUNDATIONS FOR GOOD ESTIMATIONS.

- SECOND I NEED TO INCORPORATE/CALCULATE NEW VARIABLES THAT MIGHT BE USEFUL FOR THE PURPOSE (PREVIOUS WINS/LOSSES IN A ROW, RELEVANCE OF THE GAME FOR EACH TEAM, DAYS FROM THE INMEDIATELY PREVIOUS GAME, ETC) AND IDENTIFY THE KEY ONES.

- EVERY TIME I GOT TO THE MODELLING PART I WILL LIKELY TO GO BACK TO THE DATABASE PART TO CLEAN, SIMPLIFY OR ENHANCE THE DATABASE SINCE THE RESULTS WILL GIVE ME SOME CLUES ABOUT SOME ISSUES/ERROR OR OPPORTUNITIES.

IMPORTANT

METHODOLOGY ? (II)

BEST MODEL(S) TO PREDICT RESULTS

- ONCE I CONSIDER MY DATASET IS GOOD ENOUGH I WILL WORK ON DIFFERENT SUPERVISED LEARNING MODELS TO DETERMINATE WHICH ONE ID THE BEST FOR OUR PURPOSE.

- DUE TO THE NATURE OF THE PROBLEM I WILL HAVE TO CONSIDER RECALL FOR GAMES WITH LOW-PAID QUOTAS AND TO IMPROVE PRECISION FOR THE BETTER-PAID ONES*

- I WILL DEFINE IN PYTHON A ' BETTING GAME ' TO VISUALIZE HOW MY MODEL WORKS FOR 'N GAMES' BASED ON DIFFERENT RISK PROFILES (CONSERVATIVE, HUNTER-BARGAINS AND MIXED) [EXAMPLE : APPENDIX]

- I WILL CREATE A TABLEAU INTERFACE TO VISUALIZE THE PREDICTIONS, THE MAIN VARIABLES THOSE PREDICTIONS ARE BASED ON AND THE QUOTAS.

- FINALLY, I WILL ADD A SELECTOR WHERE THE USER WILL PICK THE MOST SUITABLE OPTION(S) FOR HIM/HER BASED ON THE 3 BEFORE MENTIONED PROFILES.

* That only affects to the recommendation not to the model:

For quotas <2: odd are bigger but gains are lower = I will need to be more precise = I will need a good recall

For quotas >2: odds are lower but gains higher = 51% hit rate will be enough = I will need a good precision

WARNINGS

POTENTIAL ISSUES

- **CALCULATION OF THE FAIRE QUOTA**: IT'S INTUITIVE THAT IF THE PREDICTION TO GET A GIVEN RESULT IS 50% THE QUOTA MUST BE 2. BUT, HOW CALCULATE OTHER QUOTAS BASED ON HIGHER OR LOWER PERCENTAGES ← FOUND

- **CALCULATION OF THE IMPORTANCE OF GAME** FOR BOTH SQUADS: HOW CAN I CALCULATE THAT FCB LOST AGAINST OSASUNA ' BECAUSE BARÇA ALREADY WON THE LEAGUE AND OSASUNA WAS STILL FIGHTING TO AVOID RELEGATION?

- **IMPORTANT VARIABLES WILL BE MISSING**: RELEVANT INJURED PLAYERS, FIFA VIRUS, ETC.

- **WHEN AN EXCEEDED QUOTA IS WORTH ?** IF MY CALCULATION STATES THAT THE ODDS OF CADIZ WINNING AGAINST ATLETICO MADRID ARE RELATED TO A QUOTA OF 6 BUT THE REAL ONE IS 14, WHEN AND HOW MUCH SHOULD I BET ON THIS 'BARGAIN'?

SPRINT PLAN

PLANNING & MILESTONES

- **WEEK 1 (1st March)**

PRESENSATION OF THE PROJECT AND PROJECT APPROVAL.

- **WEEK 3 (15th March)**

LINK TO THE REPORT & FIRST MVP

- **WEEK 5 (31th March)**

FIRST LIMITED VERSION (ONLY 1 RESULT AND 1 SCENARIO)

- **WEEK 6-10 (15th May)**

DEVELOPMENT PHASE 1 & GATHERING OF ISSUES

- **WEEK 11 (22th May)**

SECOND TUTORIAL SESSION

- **WEEK 12-16 (30th June)**

DEVELOPMENT PHASE 2

- **WEEK 17-18 (15th July)**

IMPLEMENTATION & IMPROVEMENT

- **WEEK 19 (22th July)**

DELIVERY

- **WEEK 20 (29th July)**

PRESENTATION