

## Predictors of Catan

### Project Overview:

The goal of this project is to create a program, using machine learning, that will predict the winner of a game of *Settlers of Catan* at any given state. As input, the program will consider the resources currently in each players' hand, the strength of their placement, the number of cities each player has in play, the number of points currently held by each player, and their probability of obtaining resources in the near future. The prediction will be based on historical data that we have collected from two sources: an existing game generator found on GitHub at sorinMD/MCTS, and a game generator we will build ourselves. The existing code taken from GitHub is a project called *Smart Settlers*. This project aimed to determine if seating placement had any effect on a player's ability to win. As part of their research, they built an AI capable of playing a full game of *Settlers of Catan* with four players and this is the only part we are making use of in our project. We will compare the data gathered from this project to the data collected from our model, and using Tensorflow train and test an algorithm. The technologies we will be using will be Python3, for Tensorflow, and Java, for both game generators and for data collection.

If you are unfamiliar with the rules of *Settlers of Catan*, you can find the full rule book at: [https://www.catan.com/en/download/?SoC\\_rv\\_Rules\\_091907.pdf](https://www.catan.com/en/download/?SoC_rv_Rules_091907.pdf).

### **(20pts Total)** Pre-Existing Game Generator - McKenzie

- (10pts) Using game generator found on Github: *Smart Settlers*, review code to understand how the game works.
- (5pts) Get *Smart Settlers* to interact with our data collection methods i.e. modify/create code to output necessary data.
- (5pts) Generate and run 10,000 games and extract data.

### **(65pts Total)** Building Game Generator - Abby + McKenzie

- (5pts) Create board object
  - (3pts) 19 resource hex objects
    - Type
      - 18 are a resource, 1 is the desert
    - Value
      - Number 2-12
    - Probability
      - The likelihood of rolling that value

- (2pts) 9 port objects
  - Type
    - 1 grain
    - 1 lumber
    - 1 ore
    - 1 livestock
    - 1 brick
    - 4 non-specific
  - Trade Ratio
    - 5 2:1
    - 4 3:1
- (5pts) Create resource decks - 19 cards each
  - Lumber deck
  - Livestock deck
  - Grain deck
  - Brick deck
  - Ore deck
- (5pts) Create development deck
  - 2 “monopoly”
  - 2 “year of plenty”
  - 2 “road building”
  - 5 “victory point”
  - 14 “knights”
- (30pts) Create a player object
  - (2pts) Victory point count
  - (2pts) 15 road objects - empty object, a simple place holder
  - (2pts) 4 city objects - keep track of their placement on the board
  - (2pts) 5 settlement objects - keep track of their placement on the board
  - (2pts) Deck/hand object which is comprised of resource/dev objects
  - (5pts) Player understand rules of game and legal moves
  - (15pts) Decision making
    - (5pts) Resource use
    - (4pts) Development use
    - (6pts) Basic trading
      - Trading is 4:1 to the board or with any port object the player has access to.
- (2pts) Create robber object
  - Understand only its position on the board
- (12pts) Gameplay

- (4pts) Initial set up
  - Randomize resource location, values will be the same each time
- (2pts) Randomly decide which player goes first
- (2pts) Initial placement is also random
- (4pts) Resource allocation
  - (2pts) Allocate starting resources
  - (2pts) Allocate resources (if any) to each player on each die roll
- (5pts) At the start of each round collect data from each player (see Data Collection)
- (1pts) Run 10,000 games

**(15pts Total)** Data Collection - Abby

- (8pts) Scripts will be written to collect data from each player at each round in the game
  - Placement strength
  - Hand strength
  - Allocation Probability
  - Number of cities in play
  - VP total
- (3pts) Collect data
  - One sheet for data from existing generator
  - One sheet for data from our generator
  - One sheet comprised of the combined data
- (2pts) General overview of data to remove outliers
- (2pts) Divide data into training set and testing set

**(40pts Total)** Machine Learning Prediction - McKenzie + Abby

- (15pts) Tensorflow research
- (15pts) Write prediction algorithm
- (3pts) Make predictions based on each generator individually
- (4pts) Make predictions based on both generators
- (3pts) Make comparisons about accuracy

Grading Scale:

**150pts Possible**

$\geq 135\text{pts}$	A
134pts - 120pts	B
119pts - 100pts	C
$\leq 99\text{pts}$	D