FirstWork AI and Database Project

Dianne Catapang, Christian Gideon, Dillan Lopez, Andrea Ramirez
Sponsors: Thaddeus Steele (‘20) and Patrick Faga (‘20)
Faculty Mentor: Dr. Jay Kunin

University of San Diego
SHILEY-MARCOS SCHOOL OF ENGINEERING

FirstWork Introduction

- FirstWork app helps children with autism develop valuable speech-learning skills
- Blocks entertainment apps until a session set by a parent or therapist is completed
- Children will learn at their own pace
- Parents and therapists need help creating lesson plans for their child’s learning needs

Problem Statement

Develop an algorithm that takes a student’s lesson progress based on how many questions and what kind of questions they have completed successfully, and generate a optimal and suitable lesson plan for the child.

Project Development

- Created an algorithm whose input is a csv data file where each point is a question that the student answered
- Preprocess data so that all the data is within a scale of 0 to 1 to make it easier to cluster
- Determines the proper amount of clusters in algorithm using silhouette score method
- Clusters data into groups that share similar attributes so that we can examine which attributes are positively affecting the learners score

Clustering Algorithm

The approach: Agglomerative Hierarchical Clustering. Input is completed questions.

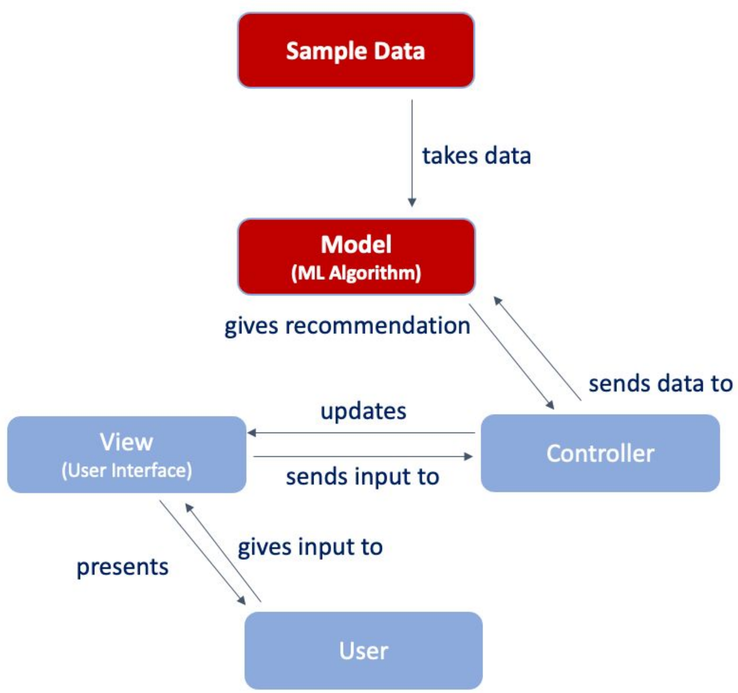
Parameters:

- Time of Day
- Reward Time
- Shuffled vs. non-shuffled
- Questions per Lesson
- Lessons per Session
- Number of Attempts (at a question)

How do we make our unique recommendations?

- Number of Attempts -> Correctness (1/attempts)
- Sort data points into clusters - similar data gets put into the same clusters
- Look for the cluster with the highest overall correctness; recommendations are the averages for each parameter in that cluster

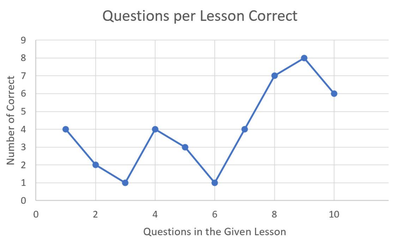
Software Architecture

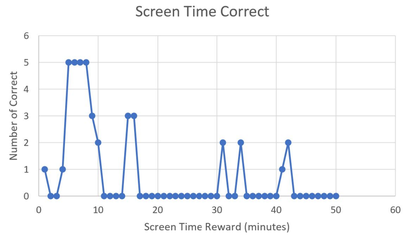


Accomplishments

- Successfully implemented clustering algorithm to output recommendations that are unique to each user
- Created documentation for project so it is available for future developers

Sample input data:





Terminal output:

```
Number of clusters: 27
Recommended Time of Day: 8:00am - 12:00pm
Recommended Reward Time: 6
Recommended Shuffle: shuffled
Recommended Questions Per Lesson: 9
Recommended Lessons Per Session: 8
```

In the future, the FirstWork team can sync the algorithm to their application, perform the clustering algorithm on real user data, and output the recommendation to the user.

Technical Challenges

- Learning new tools (Python pandas, scikit learn, and others)
- Adjusting to agile methodology (Jira)
- Research on neural networks vs. clustering
- Using randomly generated data vs real data

Acknowledgements

- Thank you Thaddeus and Patrick from FirstWork for sponsoring our project.
- Thank you to our mentors Jay Kunin. Chuck Pateros and Chuck Wolfe.