Sibyl

Team: Excelsior
Jake Yim, Skylar Winig,
Sahil Naik, Rutvik Jha,
Robin Wang
Problem

Rental Property Managers:

- They have a lot of accounting data sitting in databases with valuable insights into their properties’ financial and operational health
- They don’t have the time and expertise to laboriously go through the data
- The only current projections are through tabular reports which are often overlooked
- Having visual financial projections would help them manage expectations and optimize operations
Solution

- Create machine learning models to forecast property management expenditures, income, and other accounting categories
- Train our model using the vast plethora of user-specific property management data
- Leverage dashboard to present predictions using dynamic and interactive graphs that reveal clear insights into accounting health
Functionality

Feed in Data

Property Managers
Input accounting data and make selections of pretrained models to see visualized projects

SIBYL

Machine Learning Developers
Stores pretrained ML/DL models. Utilizes tools such as PyTorch, Scikit, Tensorflow

Model Export
Demo

http://localhost:3000/
Backup Demo
Technologies

- Containerization
- Frontend
- Backend
- ML
Novelty

01 Tabular Reports
Transform Appfolio's tabular reports into something more meaningful for property managers who may not have data science experience.

02 Data Modeling
Process data with graphs and visualization techniques to show real model training in a more insight-friendly environment for property managers.

03 Predictive Analysis
Display model inference and provide accessibility to various models to allow for more complex and thorough predictions.
Technicalities

Data/Feature Analysis

- Visualize data to look for trends, outliers, bounds
- Infer meaningful features from categories

Example:
- Use ordinality to transform dates into a feature that is trainable by a model

Prototyping Architecture

- Construct model architecture with an understanding of the characteristics of input features and how to best map them to the output

Example:
- Produce a Regression, Random Forest, DL, etc

Training/Inference

- Train architecture on training dataset
- Evaluate performance on test set or inference dataset

Example:
- Evaluate the results and error of the model, iterate hyperparameters, and extract better features
Model Development SW Infrastructure

- Docker for containerized environments to run ML models
- Created open ended scripts, with documentation, to run inference on model prototypes for simpler interface for users/front end team

Where STARTTIME and ENDTIME are in format YYYYMMDD and ID is an integer id corresponding to a GL Account ID.

```
python3 scripts/model.py --time STARTTIME,ENDTIME --id ID
```
Challenges

**Machine Learning**
- Difficult to discern meaningful features from the dataset for future projections
- Importance of Pre-Processing
  - handling empty and null fields -> filtering for valid data points
  - bolster features instead of suppress: normalization, standardization, etc

**Frontend**
- ML model agnostic data visualization
  - Ability to take in data of different types and predictions from different models with different formats
- UI/UX design
  - Data range and density
  - Interaction design
Next Steps

ML/Backend:

- Work with AppFolio to learn better feature extraction or find datasets with better features
- Transition to backend development
  - Build API infrastructure to support model export and calls from front end
  - Focus on creating support for various formats of models, instead of prototyping model

Frontend:

- Dynamic sliding window for graphs to filter by date (year, month, week)
- Customize dashboard with draggable graph components
- Search for past predictions
Thank you