

AUCU Wait Time Predictor

A Crystal Ball Fit For Deployment

David Dodds
Sept. 10, 2021



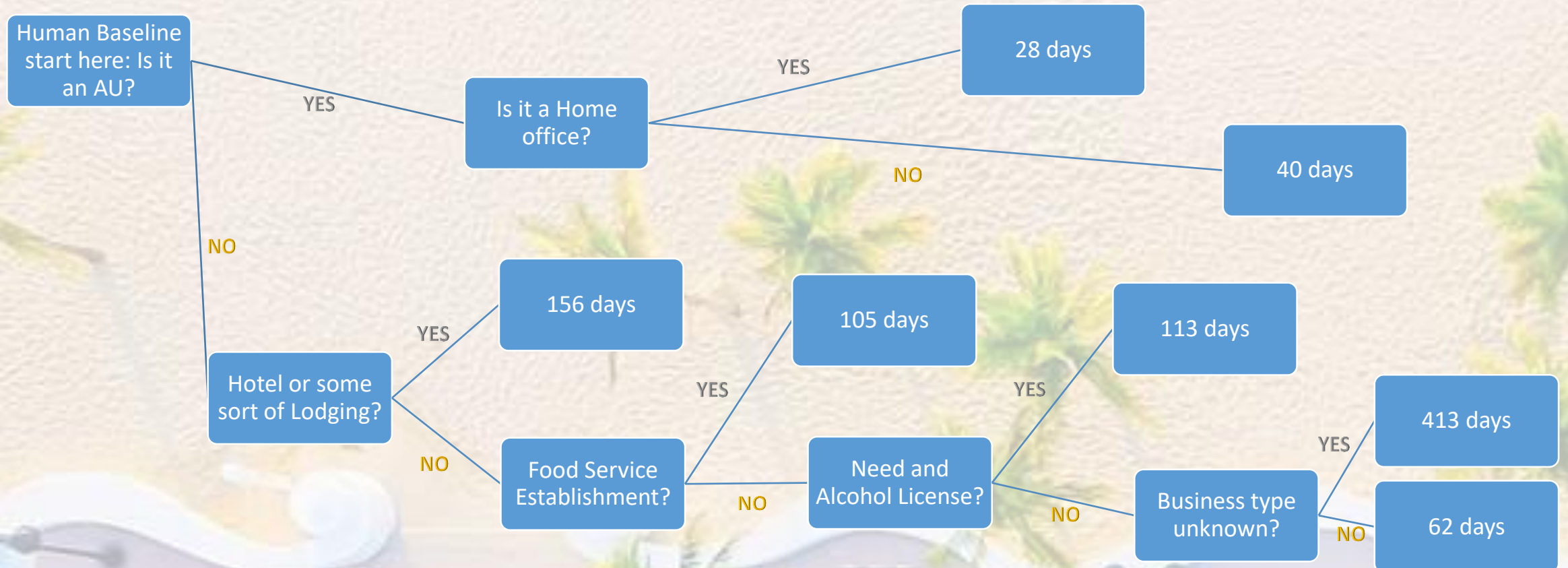
Demo of the Predictor



Baselines

“90 Day Promise”: tells all applicants they can expect their license in 90 days.

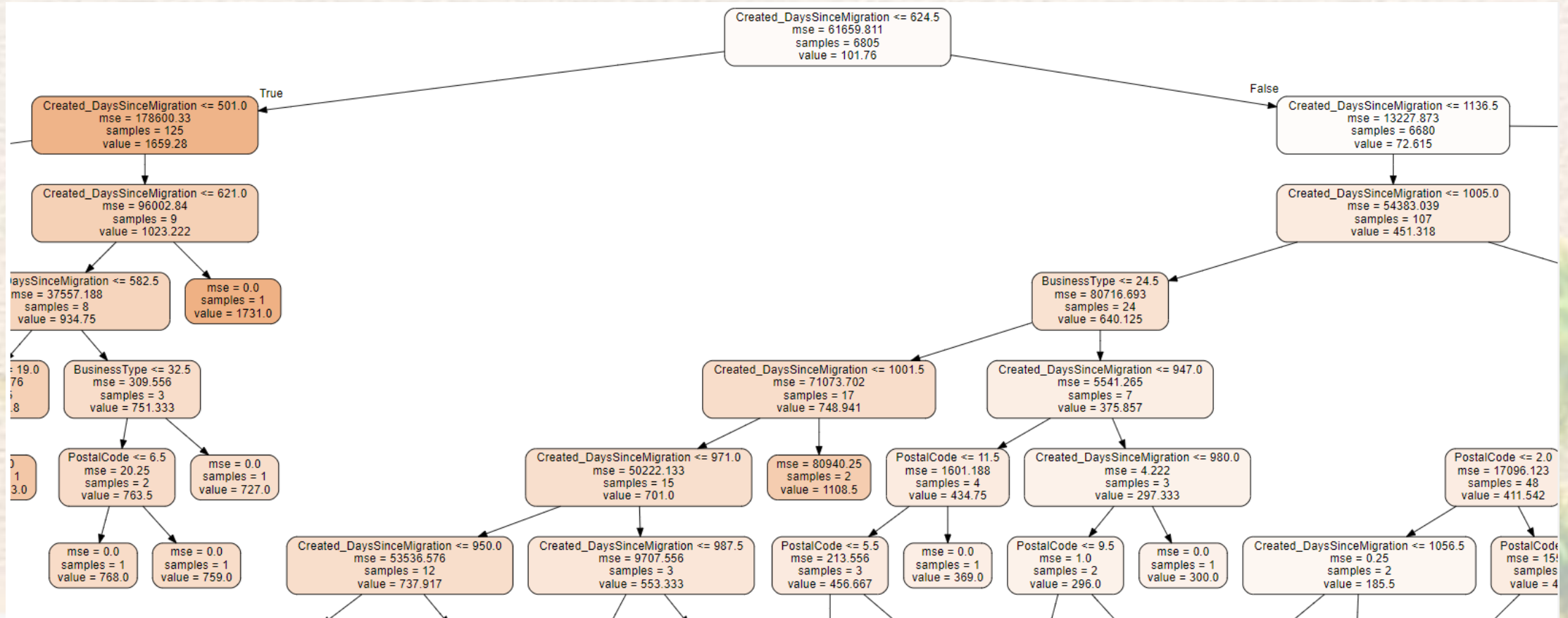
“Human Baseline”: using a flowchart like approach, specifies a wait time prediction based off of features of the applicant, and what we know about which types of applications take longer.



Baseline Error Scores

	90 Day Promise	Human Baseline
Mean Absolute Error (MAE) How any days off was an average prediction?	104 days	106 days
Root Mean Squared Error (RMSE) Like MAE, but punishes predictions that were way off.	253 days	234 days
R^2 Score Measure of how correlated the predictions are with the actual wait times of past data.	-0.24 %	14.56%

MVP model: Decision Tree



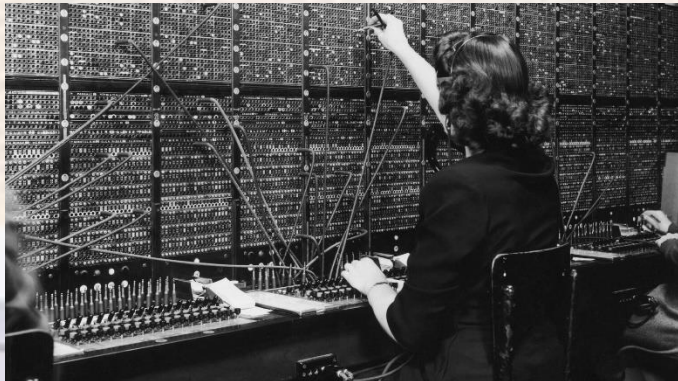
MVP model: Error Scores

	90 Day Promise	Human Baseline	MVP Decision Tree
Mean Absolute Error (MAE) How any days off was an average prediction?	104 days	106 days	71 days
Root Mean Squared Error (RMSE) Like MAE, but punishes predictions that were way off.	253 days	234 days	150 days

9 models later

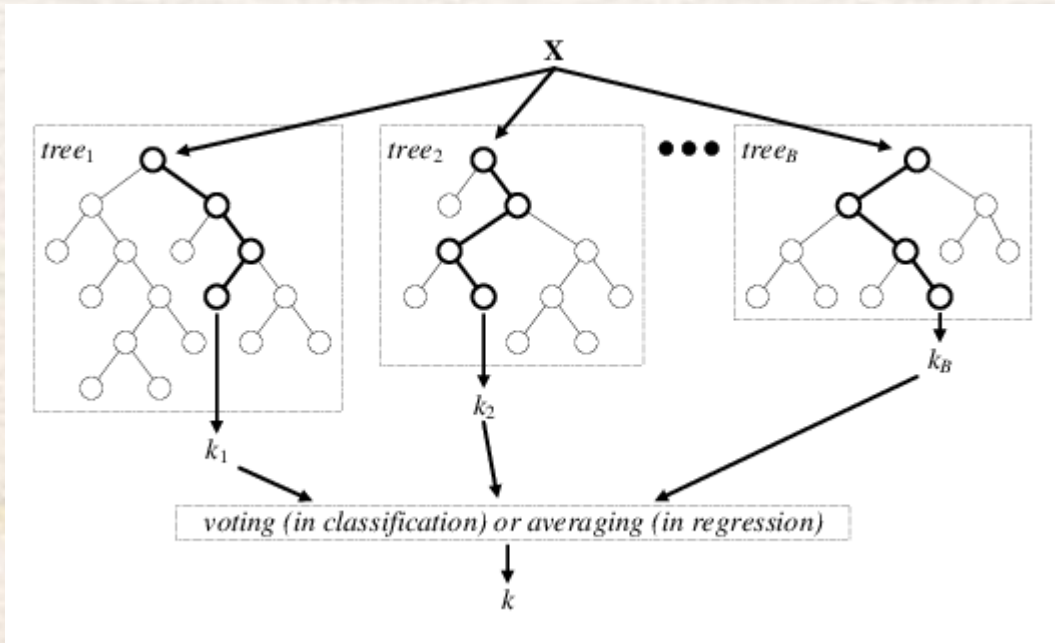
	Human Baseline	MVP Decision Tree	9...	Mach 9 Random Forests
Mean Absolute Error (MAE) How any days off was an average prediction?	106 days	71 days	models...	25 days
Root Mean Squared Error (RMSE) Like MAE, but punishes predictions that were way off.	234 days	150 days	later...	29 days

What tweaks were made?



- Stripped out the creation year as a decision making factor for the model
- Iterated optimal timeframes to use
- Excluded old data (28 months was the cutoff for this model)
- Tuned hyperparameters
- Target Encoded the data
- Used Random Forest model instead of a single Decision Tree

What is a Random Forest anyway?



- Many trees created (trained) on random subsamples of the data
- New inputs are handed to the “forest” by user application inputs
- Each tree votes on a predicted number of days
- Forest returns average vote (prediction) to user

Feature Importances

Feature	Importance Index
Business Type	66.82 %
Day of the Year of Application Creation	12.42 %
Day of the Month	8.56 %
The Postal Code	8.26 %
Day of the Week	3.03 %
Month of Application Creation	0.91 %

Extracting full value from this predictor

Rough estimates assuming 1% of applications are directly effected

Monetary Juncture (opportunity)	Making the most of it	Conservative yearly savings estimate
Entrepreneur's Planning / Launching Process	Direct entrepreneurs to use predictive tool to plan their timelines	~\$200,000
Application / Inspection / BTR Fee Revenue	Direct applicants to use predictive tool to inform their decision on whether they are ready to begin an application	~\$20,000
Smoothing Zoning's application process	Use predictive tool to manage applicants expectations while in process, use insights to inform process improvements	~\$5,500
Tax Revenue	Use predictive tool to smooth business opening, reduce failure rates	~\$80,000

Toward business X neighborhood affinity

- Maria's vision: an online tool entrepreneurs can use to forecast where there business is most needed and most likely to succeed
- Reductions in storefront vacancy and crime
- Improved Neighborhood social capital and residency rates
- Improved educational outcomes and employment opportunities

Next Steps

Deploy the model online

- Pickle the model
- Stitch the model into Streamlit or Google Environment Site

Pipeline to ensure models can adjust as new data flows in.

- Clean new data
- Explore new data
- Automate new model iteration