Predicting Failure to Appear (FTA) in court cases

Ping Tu, Olatunji Akinbule, Zhijie Ren

Data Analytics, George Washington University

ptu16@gwmail.gwu.edu

akinbule_ola@gwmail.gwu.edu

zren88@gwmail.gwu.edu

***Abstract*— Charges of failure to appear (FTA) are common across both court and arrest records. FTA was the top felony arrest charge from FY 2015 and FY 2017, making up nearly 20% of felony arrests, and was the fifth most common offense in court records between 2010 and 2016. While there might be a multiplicity of reasons for an FTA, Focus groups and interviews with Public Defender Services (PDS) and Pretrial Services (PSA) staff suggest that lack of transportation is a significant factor in failures to appear. Some defendants may be indigent or may prioritize other needs like food or rent above transportation. Additionally, some defendants already receive metro cards from their attorney with a positive effect on their appearance by their attorneys. Given that transportation is reported as a barrier to court appearances, our objective is to develop machine learning models to predict the likelihood of the appearance of a defendant using data that is available at the outset of a case. This will allow for greater precision through a stratified random controlled trial in assessing the treatment effects associated with receiving a free MetroCard at the time of their release..**

Keywords*— *Machine Learning, Featuretools, Failure to Appear (FTA)

I. Introduction

Charges of failure to appear (FTA) are common across both court and arrest records. FTA was the top felony arrest charge from FY 2015 and FY 2017, making up nearly 20% of felony arrests, and was the fifth most common offense in court records between 2010 and 2016.

Missing a court-mandated appointment, especially a hearing before a judge or other judicial entity, can have serious negative consequences for defendants and their families. Defendants who fail to appear may have additional charges filed against them, may have bench warrants issued against them, may have their bail revoked, or may be detained. As indicated by extensive research over the last three decades, many of these costs are borne not only by defendants, but by their families as well.

Focus groups and interviews with PDS and PSA staff suggest that lack of transportation is a significant factor in failures to appear. There are many reasons why transportation is a likely barrier to appearance. Defendants may be indigent; they may lack the funds on the day of the appointment because their cash on hand fluctuates; or they may prioritize other needs like food or rent above transportation.

Anecdotally, some PDS clients already receive metro cards from their attorney, and the attorney has described the effect on attendance as positive. Given that transportation is widely reported as a barrier to court appearances, our objective is to predict the likelihood of the appearance of a defendant based on past data. While there might be a multiplicity of reasons for an FTA, the proposed project will attempt to identify defendants with a high risk of FTA and test whether the distribution of prepaid metro-transit cards can increase court appearance rates.

We want to develop Machine Learning models that will allow us to assess the likelihood that a defendant will fail to appear using only the data that is available at the outset of a case. This will allow for greater precision in assessing the treatment effects associated with receiving a free MetroCard at the time of their release.

II. Data

The theoretical population of this project is all defendants over the course of the existence of the DCSC including defendants with published and unpublished court records. Some court records are not published because they might involve juveniles, are unavailable at the point of our data scraping, sealed or expunged. And the accessible population is all case information published to the DC Superior Court website. The sampling frame for this project is the entire set of scraped misdemeanor court case information from 2006 through 2019. The data source of this project is https://eaccess.dccourts.gov/eaccess/search.page. A major challenge with writing a web crawler on the DC Superior Court Website was getting around the captcha, however we used cookies to harness a longer valid session and dramatically increase the data extraction efficiency although our crawler is not fully automated. We scrape two kinds of case from this website, one is felony case and the other one is misdemeanor case.

They have different structures in detail page. Felony has event table whereas misdemeanor has docket as equivalent and financial docket as a simplified docket. We have 29268 Felony Crime cases, including CF2 and CF3, and 103967 Misdemeanor cases, 31.04% of which are labeled FTA (32273). For now, we only focus on Misdemeanor cases, because we have more data for them.

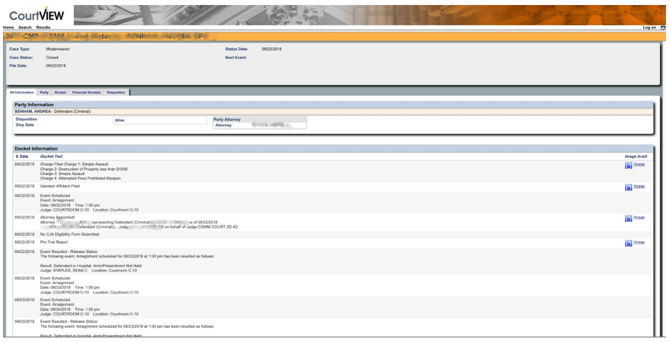

Fig. 1 DCSC CourtView

A. Term Definition

Case:

Represents each unique legal case

Party:

The defendant or consumer who is the subject of a court case

Court Docket:

Record of the proceedings of a court case usually containing information about the parties such as the judge, defendants and case events. See Fig. 1.

Receipt:

Payment records that may detail fines and fees paid by defendants (parties) to the DC superior courts for bench warrant bonds or to the Crime Victims Compensation Program that provides financial assistance and reimbursement to victims of crime.

Attorney:

Member of the judicial profession who represents a client in court when pleading or defending a case.

Bench Warrant:

Warrant issued by a judge for the arrest of a person who has violated court rules and is in contempt of court. The warrant gives police the powers to apprehend the defendant and bring them before the court to answer contempt charges.

Failed to Appear:

When a defendant obligated to show up in court fails to appear and also has no representation.

B. Court Proceedings / Docket Events

Court proceedings can entail a number of unique events including arraignment, status, admission, review, bond-review, probation show cause and sanction hearings over the duration of a case.

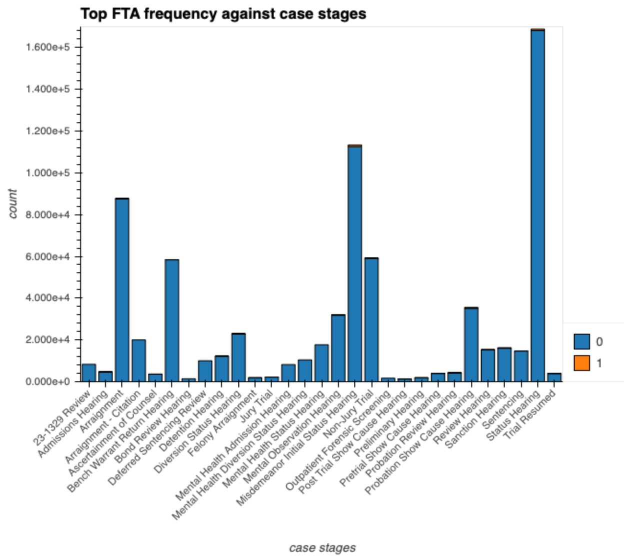

Fig. 2 top FTA frequency against case stages

Fig. 2 above shows the almost insignificant FTAs in orange at different court events as we sought to explore if certain court proceedings was more prone to FTAs.

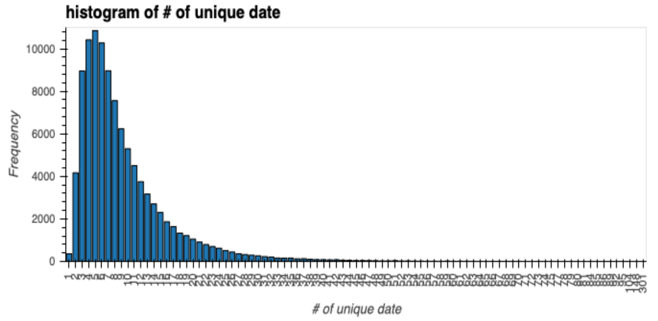

Fig.3 Histogram of number of unique date

Defendants are mandated to be present at their court proceeding, however represented defendants can have their attorneys show up for some of their court event cases on their behalf.

There's a high burden for people who are not represented. Indigent defendants are more likely not to have hired an attorney and may have to depend on the public defender services.

Failure to Appear is characterized by a situation where a defendant and their representative is absent for a court proceeding, a bench warrant for their arrest is usually issued by the court.

Fig. 3 shows the long tail distribution of the number of unique court event cases over the lifetime of a case for which a defendant or their representative is required to appear in court. The range from the mode standing at 5 appearances (10850 cases) and a maximum of 301 appearances depicts the region with a marked increased burden of appearances. The number of appearances required in a case might be contingent on the complexity of the case, some extenuating circumstances and or failure to appear. As the number of court events for each case increases it is easy to project that there is an increased risk in the likelihood to fail to appear. While our overall objective is to assess the impact of a metrocard intervention especially for indigent defendants, we acknowledge that there is a deluge of other reasons a defendant might fail to appear including mental health challenges, hospitalization, ignorance about the severity of missing court appointments and forgetfulness or confusion about scheduled court events. For example the case with the 301 appearances was reported in the court docket as "Defendant in Hospital" never making it to any arraignment proceeding but with an attorney present at each court event till their case is quashed after a year with no FTA on record.

III. Exploratory Data Analysis

From Fig.4 below , the graph below shows the rate of FTA from 2016 through to 2019 at the DC Superior court with an average of about 30% FTAs across the years.

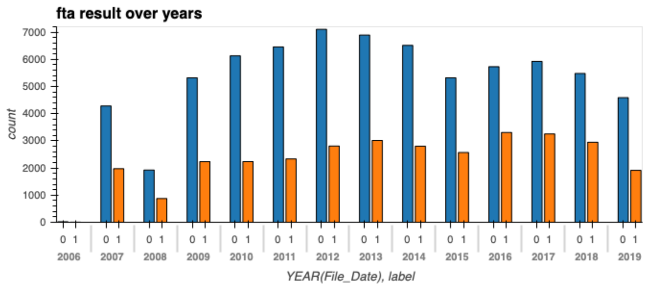

Fig.4 FTA result over years

In addition, the number of judges shown in Fig 5 shows the distribution of the number of judges involved in particular cases. The mode for the number of judges is 3 judges (41271 cases) deliberating on each case and a mean of approximately 4 judges in our right skewed distribution below.

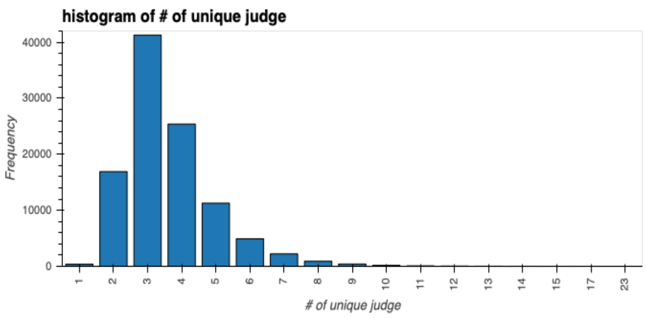

Fig.5 Histogram of number of unique judge

The top 10 charges recorded are simple assault, bail reform act, destruction of property, failure to appear, possession of a controlled substance , sexual solicitation, shoplifting, theft(second degree), threats to do bodily harm and unlawful entry. Fig 6 shows the proportion of the top 10 charge types and the number of court appearances. Fig 7 The normalized diagram shows the proportion of possession crimes increases with the number of court appearances.

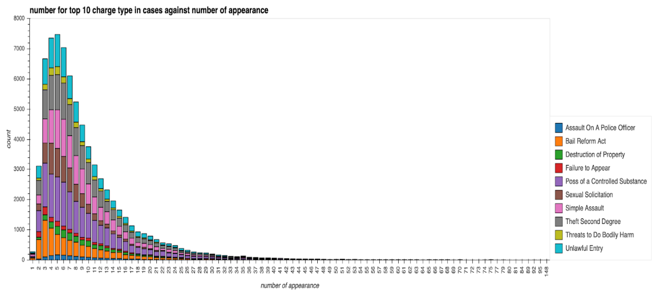

Fig.6 stacked barchart for top 10 charge type in cases against number of appearances

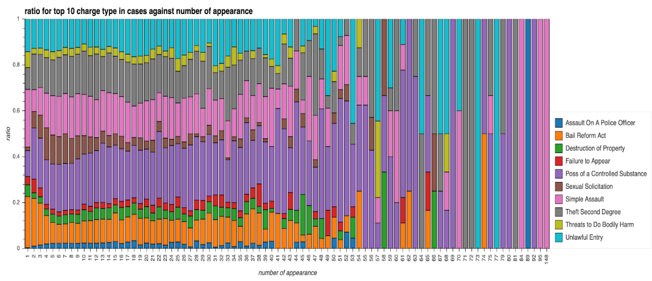

Fig.7 Normalized into 100% (based on Fig.6)

IV. Methodology

A. Feature Engineering

Featuretools is a python package that we employed to automate our feature engineering process. There are three main entities that needs to be created, which are party(consumer/defendant), cases and dockets. See Fig. 8. Cutoff time is set using the File Date of cases to prevent data leakage because we only allow Feature Tools to generate the features for current case using the data before it. But Featuretools also includes the data that happens right on cutoff time.

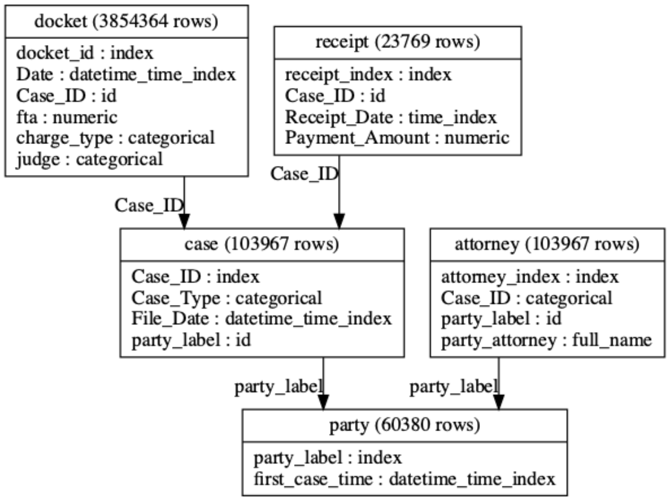

Fig. 8 ERD

For example, to get previous FTA count for current party, we need to subtract the party.SUM(receipt.Payment_Amount) by SUM(receipt.Payment_Amount). Fig. 9. proves that our assumed feature is important. Because it says for people who don't have a fta history before, it is very likely that ¾ of them will show up on the court. And once they have at least fta in their history records, the chance whether they will show up is about 50%. This makes sense because defendant doesn't not know what the consequence of FTA is and they tend to obey.

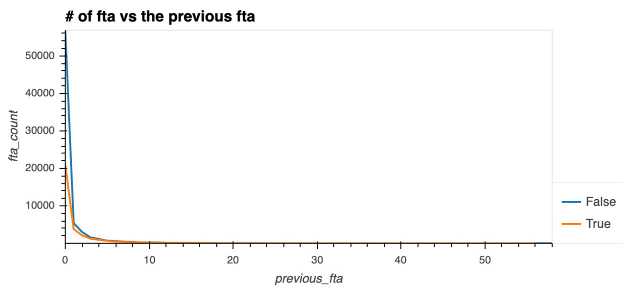

Fig.9 number of fta against previous fta

Finally, the features we get from Featuretools are listed as follows:

```
Index(['label', 'COUNT(docket)', 'AVG_TIME_BETWEEN(docket.Date)',
       'TIME_SINCE_LAST(docket.Date)', 'TIME_SINCE_FIRST(docket.Date)',
       'NUM_UNIQUE(docket.charge_type)', 'NUM_UNIQUE(docket.judge)',
       'COUNT(receipt)', 'SUM(receipt.Payment_Amount)',
       'CUM_COUNT(party_label)', 'YEAR(File_Date)', 'MONTH(File_Date)',
       'SUM(docket.CUM_COUNT(Case_ID))', 'NUM_UNIQUE(docket.YEAR(Date))',
       'NUM_UNIQUE(docket.MONTH(Date))', 'SUM(receipt.CUM_COUNT(Case_ID))',
       'party.COUNT(case)', 'party.AVG_TIME_BETWEEN(case.File_Date)',
       'party.TIME_SINCE_LAST(case.File_Date)',
       'party.TIME_SINCE_FIRST(case.File_Date)',
       'party.NUM_UNIQUE(case.Case_Type)', 'party.COUNT(docket)',
       'party.AVG_TIME_BETWEEN(docket.Date)',
       'party.TIME_SINCE_LAST(docket.Date)',
       'party.TIME_SINCE_FIRST(docket.Date)',
       'party.NUM_UNIQUE(docket.charge_type)',
       'party.NUM_UNIQUE(docket.judge)', 'party.COUNT(receipt)',
       'party.SUM(receipt.Payment_Amount)', 'party.COUNT(attorney)',
       'party.NUM_UNIQUE(attorney.Case_ID)', 'party.YEAR(first_case_time)',
       'party.MONTH(first_case_time)', 'previous_receipt', 'previous_fta'],
      dtype='object')
```

Fig. 10 Features

V. Evaluation

TABLE I

Confusion Matrix

		Predicted Values	
		"FTA"	"Non-FTA"
Actual Values	FTA	True Positive (TP)	False Negative (FN)
	Non-FTA	False Positive (FP)	True Negative (TN)

A. Recall (Sensitivity)

Recall describes the ratio of correctly classified positive FTAs against all FTAs that actually occurred. In other words, when a defendant fails to appear (FTA) how often does our model predict the failure or the percentage of all cases where the defendant failed to appear and our model is able to correctly identify. Recall is affected by the number of false negative errors.

B. Precision (affected by number of positive errors)

Precision is the ratio of the accurately predicted FTAs to the total number of predicted FTAs. Of all the predicted positive FTAs, how many were correctly predicted. Precision is affected by the number of false positive errors.

Precision shows how certain we are about an FTA and recall depicts the metric that we are not missing any FTA cases.

As a first consideration the team leaned more towards recall as the primary accuracy metric since false negatives are less tolerable.

The overall objective is to make sure we are not missing FTA cases even at the risk of incurring some overhead of false positives. The consequence of this choice in terms of the metro card policy is to provide the most metro cards to the defendants that are more likely to FTA reducing our false negatives.

Consequently a potential accuracy metric to consider will be the F2 score (□ =2).

This approach will weight recall higher than precision since for the FTA project we have determined that it is more important to correctly classify as many FTAs as possible ,even so, as FTAs are the minority class (more defendants show up for their court cases than not). From the data in 2007 - 2019 there were 10,867 FTAs (6739 unique cases) in the approximately 32,300 unique cases a 20% FTA rate.

V. Ensuring Predictability

In order to verify the predictability of our models we have earmarked court cases in 2019 as our hold out data. In addition we will be implementing multiple machine learning models in order to make a comparative assessment of the results across the best performant model to ensure some form of conclusion validity.

VI. Results

We developed three models Xgboost, Random Forest and Light GBM. Say something about XGBoost, RF and LightGBM. Fig. 11 and Fig. 12 shows the features we trained for the model.

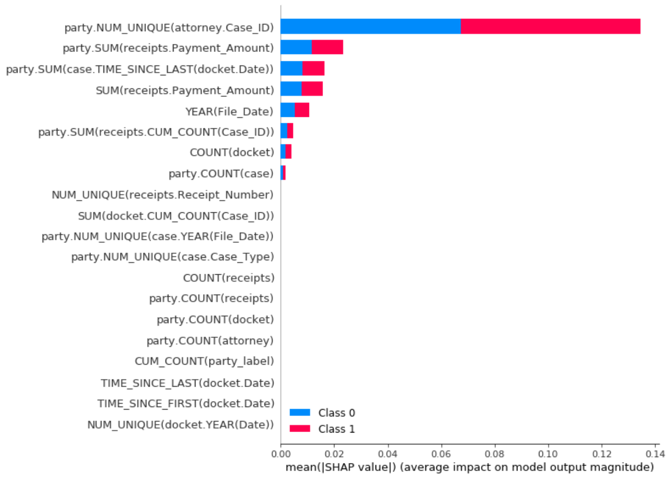

Fig.11 SHAP on Light GBM

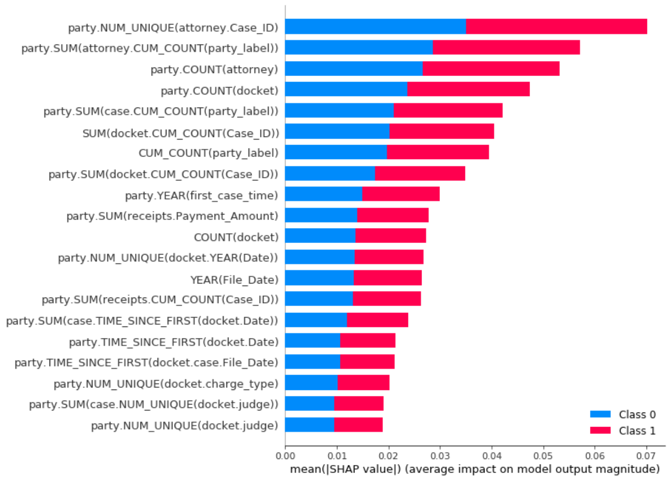

Fig.12 SHAP on Light Decision Tree

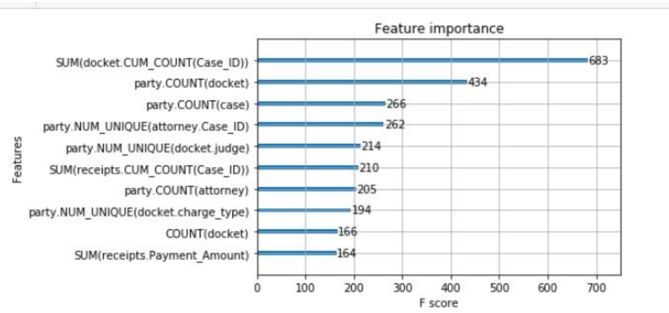

Fig.13 Feature importance

Fig.13 shows the feature importance and Fig.14 shows the visualization of the decision tree.

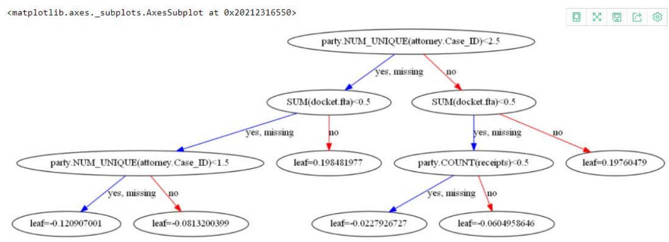

Fig.14 Decision Tree

Here is the result and Fig.15 shows the ROC curve.

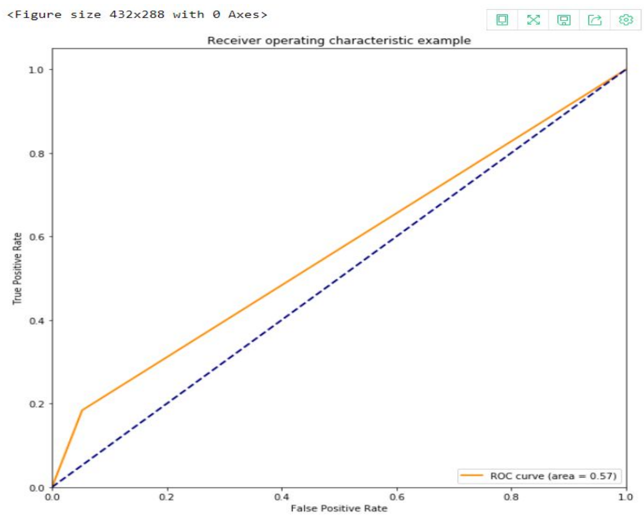

Fig.15 ROC curve

TABLE II

Recall	18.39%	When a defendant fails to appear (FTA) how often does our model predict the failure
Precision	61.79%	Of all the predicted positive FTAs, how many were correctly predicted
F1 - Score	28.3%	Evenly weighted Recall/ Precision

B. *INTERPRETING RESULTS*

Across the 3 models one feature that stands out is the **party.NUM_UNIQUE(attorney.Case_ID)** which represents the number of unique cases the attorney has handled in the past. From our research we found out that the court appointed public defender attorneys are usually overburdened with more cases than they can possibly give time to as indigent defenders are assigned to them. There is no guarantee that the public defenders would have enough time to make a representation for them especially when they do not show up. The New York Times [2] reports “On Lawyer, 194 Felony cases, and no time”. This lawyer has to do the work of 5 full time lawyers to serve all 194 clients.

Also, the **party.COUNT(attorney)** as the second most important feature in Decision Tree model. This feature represents the number of attorneys this current defendant has had in the past. This feature depicts that the more attorneys assigned to a defendant’s case the longer their case might linger and the more prone they are to FTAs

VII. Limitations

We have undertaken the task of developing a model to predict the FTA based on the limited features (past court records and appearance rates) provided by the DC Superior website court. Our model cannot account for other confounds that we cannot control for in each case including mental and physical health of defendants, socio economic factors such as education, employment, income, housing and other personal information that are potentially predictive features.

On a different note, a defendant may be recorded under multiple party names in DCSC CourtView data. Currently, we assumed each different party label represents different individual. But Entity Resolution requires more data from other data source to get features including age, date of birth. This is a process that should be added into the workflow of this project in the future.

Precision at K will be a potential second metric as we understand that the proposed Metro card intervention might be capped to a specific number. For example if there are only 1000 metro cards for distribution, we want to consider a metric that gives us the best likelihood of 1000 true positives.

VIII. Conclusion

The developed models, which have poor predictive results have shown that a key predictor to determining the FTA of a case is the number of unique cases an attorney is assigned and the number of attorneys that have been assigned to a case.

Ideally with more features that are able to represent defendant’s socioeconomic features, mental health and other potentially predictive features we would be able to obtain better predictive results.

IX. Future work

A stratified randomized control trial with a prepaid metrocard treatment is the proposed methodology to rule out confounds and determine the effectiveness of the intervention.

The Lab @ DC is partnering with the Public Defender Service to develop a pilot program designed to support court appointment attendance by defendants by providing them with free metro cards. To assess the marginal effect of providing defendants in criminal cases, The Lab will assign participants to a treatment (metro card) arm or control (no metro card) arm.

The design will involve assigning defendants to a treatment or control condition based on a stratified random sample design. The assignment could be stratified by case type, age, etc. but the ideal form of stratification -- for which all of these other strata would be proxies -- is risk of failure to appear itself.

In other words, defendants are blocked by their FTA likelihood are randomly assigned into treatment (receiving prepaid metro cards) and control groups with no intervention to avoid the selection bias. A predefined time duration, probably a few months within which a number of court appearances are scheduled, is allowed to ascertain any effects of the intervention based on their subsequent appearance rates.

References

[1] https://www.featuretools.com/project/predict-appointment-no-show/

[2] https://www.nytimes.com/interactive/2019/01/31/us/public-defender-case-loads.html