GROUP RESEARCH PROJECT & PAPER

Artificial Intelligence for FinTech
CAP5619-24Spring 0V61
University of Central Florida

Group 4
Vincent Bortone
Andrew Kasztan
Thomas Harshman
Max Laguerre
Luis Mejia

Teachers & Tas
Ramya Akula
Sourav Joshi

April 24, 2024

Keywords: risk governance, compliance, sales and use tax audit, AI, machine learning

INTRODUCTION

Our present research project focuses on risk governance. More particularly, it identifies and assesses a substantial risk that enterprises often overlook: Sales and Use Tax law compliance. Yet, non-compliance with sales and use tax laws and regulations can have dire impact on enterprise functioning and even survival. It is a risk too severe for enterprises to ignore. Our research aims at providing a tool for enterprises to identify, monitor, and mitigate such areas of risk.

1. DEFINITION OF THE PROBLEM

For entrepreneurs and businesses, compliance with government regulations is often an after-thought if a thought at all. They focus on development and production, acquisition and cashflow, marketing and sales, and so forth. That is until the government knocks at their door.

In Florida, one government agency that consistently knocks on business doors is the Florida Department of Revenue (DOR), to ensure sales and use tax compliance. With its interagency state and federal data match, its state-of-the-art lead programs, and the sharpness of its auditors, DOR sales and use tax audit is mostly certain to discover non-compliance. In the other side of the diptych, businesses are too often unaware of the variety of third-party data the department have access to and collect through nightly interagency digital data interface. Then, they grow dismayed. The terror starts with the Notice of Intent to Audit Books and Records. Less than a year later, the business receives the Notice of Proposed Assessment which conveys one important message: noncompliance is costly.

This research project offers the opportunity to businesses to predict and by that to avoid or mitigate the risk of an audit. Using machine learning, the algorithm predicts the risk of a sales and use tax audit. It puts the power in the hands of businesses. By power, we mean compliance and the avoidance of costly mistakes that cause the business to incur not only additional tax due, but hefty penalties and interests, legal and professional fees, not to mention criminal investigations and proceedings. In extreme cases, an audit means the closure of the business, especially for small businesses.

2. MARKET STUDY: BY ANECDOTES

In September 2021, Automotive inc. received the Notice of Intent to Audit books and records (form DR-840) from the Florida Department of Revenue. Automotive Inc. (fictitious name used here for confidentiality) is a new car dealer which main operations consist of selling and leasing new and used cars, trucks, and SUVs. The chief Financial Officer, Mr. Leger Cramer, reviewed the document with no excitement. The company was undergoing a historical merger. So, the timing could not be worse. Thus, Mr. Cramer was confident that the audit would be quick and would conclude in the company favor, with no assessment. He maintained all along that Automotive Inc. has all the structures in place and has been complying with sales and use tax laws.

About 6 months after the audit commencement, an initial assessment of over three million dollars, plus penalties and interest, crushed Mr. Cramer’s confidence. His usual good manners turned into rage and bitterness. Then, the audit starts to drag way longer than Mr. Cramer expected. Facing the disruptive and unexpected scope of the audit, Automotive inc. hired both a Certfified Public Accountant (CPA) firm and a law firm specialized in sales and use tax audit to continue with the audit and manage the protest. Professional fees, CPA and legal typically range from 10% to 20% of the assessment. In this case, up to half a million dollars.

Being an established company, Automotive inc. survived beyond the audit. Unfortunately, this was not the case for Hollo Mercer (confidentiality protected). For over 12 years he had owned a convenience store. That is until DOR slapped him with a $280,000.00 bill as the result of a sales and use tax audit. Incapable of making the payment, he signed a stipulation agreement with the department to pay the bill along with the penalties and the accruing interests within a year, with twelve monthly installments. After a year, Mr. Mercer did not pay even a quarter of the dues. DOR initiated enforcement actions. After DOR froze the business bank accounts, Mr. Mercer could not respond to the business daily operations, like buying inventory or paying rent. He closed the door and lost everything.

Those two examples illustrate the cost of non-compliance with DOR sales and use tax laws. In the case of Automotive inc. the non-compliance was not intentional; while with Mr. Mercer there are indicators that the non-compliance might have been fraudulent. Still, in both cases, the financial and legal consequences might be devastating. It is in that context that we offer the Audit Risk AI algorithm as a technological redeemer.

3. MARKET STUDY: BY LITIGATIONS

In the previous sections we provided anecdotal evidence of the cost of noncompliance. Because the source of such anecdote is internal and to respect the taxpayer's right to confidentiality, we could not reveal the name of the subjects. Thus, public records offer us more direct evidence of such costs. Informal protest through the DOR Technical Assistance and Dispute Resolution (TADR) is the normal channel to resolve unagreed audits. When informal protest fails, the protest makes it all the way to the administrative or judicial court system. The consequences are the uninterrupted disruption of the business process along with skyrocketing professional fees and litigation costs. The table below displays seven controversies as a sample to illustrate the point.

Taxpayer	Controversy	Amount	Court
Cady Studios LLC v. DOR	Documentation of exempt sales	$352,473	4th Judicial Circuit, (Duval County), Case No.: 2021-CA-2752 Amount in Controversy:
St. Johns Ship Building, Inc. v. DOR	Timing of agency action related to chapter 120 petitions.	$1,350,785	1st DCA, Case No.: 1D23-0751
T-Mobile Resources LLC. Successor in Interest to T-Mobile Resources Corporation v. DOR	Sales and use tax refunds	$5,790,358	2nd Judicial Circuit (Leon County), Case No.: 2021-CA-206
Golden Triangle Properties, Inc. v. DOR	Assessment methodology	$1,061,949	DOAH, Case No.: 23-1663
Central Shared Services v. DOR	Documentation to support refund of sales and use tax.	$3,244,844	DOAH, Case No.: 22-2228
Cemex Construction Materials Pacific LLC successor by merger to Cemex Construction Materials LP vs. State of DOR	Validity of closing agreement	$2,859,361	2nd Judicial Circuit (Leon County), Case No.: 2017-CA-476; 1st DCA, Case No.: 1D20-3541

Source: *Florida Revenue Estimating Conference - 2023 Florida Tax Handbook*
* DOR: Florida Department of Revenue

As demonstrated in the above table, sales tax assessment can be hefty. In fact, the reported amounts relate only to issues in controversy. Audit assessments usually consist of multiple exhibits, one per issue. The taxpayer may protest only the disputed ones. For instance, in Cady Studios LLC v. Florida Department of Revenue the issue in controversy is Documentation of exempt sales, for $352,473. There may have been other assessments not protested, like unreported sales, tax rate, tax collected and not remitted, and so forth. Besides, there are the penalties and interest, plus the heavy cost of litigation.

4. MARKET STUDY: BY THE NUMBERS AND BY STATUTES

Per the Florida Department of Revenue, Office of Tax Research Collections and Distributions - Florida Dept. of Revenue - taxresearch (floridarevenue.com) - Florida statewide sales between January 2023 and February 2024 amounted to $2,245,857,837,903.39. for the same period, the department collected $63,122,450,173 in sales and use tax. This is astronomical. This covers one hundred industries and sectors, from Grocery Stores to Seafood Dealers, from Restaurants, Lunchrooms, from Furniture Stores, New and Used to Manufacturing, Processing, Mining, and from Book Stores to Motors Vehicle Dealers, Trailers, Campers.

As reported in the Florida Revenue Estimating Conference - 2023 Florida Tax Handbook Including Fiscal Impact of Potential Changes - for the fiscal year 2023-2024, Sales and use tax represent 75.2% of the state total general revenue, far surpassing contribution from other state taxes and fees, like corporate income tax (11.6%), documentary stamp tax (2,6%), or corporate filing fees (1.2%). So, sales and use tax are a huge component of the Florida government financial structure. Budgets, requisitions, and expenditures across general government and judicial, education and social services, criminal justice and transportation, all depend on it. So, it should be no surprise that the department of revenue is overly aggressive at enforcing sales and use tax laws.

Florida Statutes 212.05 gives the Florida Department of Revenue the authority to levy sales and use tax on those sales, including on retail sales of most tangible personal property, admissions, transient lodgings, commercial rentals, and motor vehicles. The statutes mandate businesses to register and to collect and remit sales and use tax to the department by filing monthly, quarterly, or annual tax returns, depending on sale volumes.

Florida statutes give the department authority to examine any business books and records for sales and use tax compliance. So, the market regulated by sales and use tax is a huge one. Whether a business is a giant monopoly like Amazon or a small mom-and-pop neighborhood convenience store, sales and use tax audit is a concern for everybody. Compliance is inescapable. Even unsuspected activities like purchases made through the internet, mail-order catalog purchases, purchases made in another country, furniture purchased from dealers located in another state, and computer equipment ordered from out-of-state vendors advertising in magazines are subject to sales and use tax.

Since July 1, 2021, the new law requires marketplace providers like Amazon, eBay, and Etsy that processes sales or payments for marketplace sellers (individuals who sells through a marketplace provider) to electronically register to collect and electronically remit sales and use tax on taxable sales they facilitate for marketplace sellers for delivery into Florida.

Because of the law, businesses in Florida, which activity or products are subject to sales and use tax, must register to collect sales tax or pay use tax, as well as Discretionary Sales Surtax, Transient Rental Taxes. From https://floridarevenue.com/, here is a partial list of business activities that require registration with the Florida Department of Revenue:

- Sales of taxable items at retail
- Repairs or alterations of tangible personal property
- Rentals, leases, or licenses to use real property (for example: commercial office space or mini warehouses)
- Rentals of short-term living accommodations (for example: motel/hotel rooms, beach houses, condominiums, timeshare resorts, vacation houses, or travel parks)
- Rentals or leases of personal property (for example: vehicles, machinery, equipment, or other goods)
- Charges for admission to any place of amusement, sport, or recreation
- Manufacturing or producing goods for retail sales.
- Selling service warranty contracts
- Operating vending machines or amusement machines and concession stands.
- Providing taxable services (for example: investigative and crime protection services, interior nonresidential cleaning services, or nonresidential pest control services)
- Out-of-state businesses selling in Florida that have any number of transactions with total sales over $100,000 in the prior calendar year.
- Marketplace providers facilitating remote sales into Florida.

5. **SCOPE OF THE PROJECT**

Now, we know with certainty that compliance is a sizable problem that covers an extensive market. Fortunately, we can use artificial intelligence as a technological breakthrough to solve the problem. What we are dealing with is a simple classification problem: compliance versus non-compliance, audit risk high versus low. Even with its simplicity, our proposed algorithm is mighty. Its value expands way beyond a yes or no prediction. It can prevent businesses from the nightmarish scenario of a sales and use tax audit. It can also become a tool in the hands of sales tax professionals, CPA, auditors, and attorneys involved in the auditing process on the clientele side. It allows them to build audit prevention infrastructures as well as defense strategies for their clients.

6. DATASET

The Florida Department of Revenue does not publish data relating to its auditing activities. Fortunately, through our research and consulting with industry insiders, we recovered the department lead algorithm. From there, we generate a simulated audit risk dataset. The dataset consists of the result of 100160 sales and use tax audits performed on different sectors and industries. Our main data source is the North American Industry Classification System (NAICS).

Per the United States Census Bureau, *the North American Industry Classification System (NAICS) is the standard used by Federal statistical agencies in classifying business establishments for the purpose of collecting, analyzing, and publishing statistical data related to the U.S. business economy*.

To construct the dataset, we use the proprietary data industry averages from NAICS. The data are proprietary (membership fee-based access) and not publicly available. Its use here is for the sole purpose of this academic exercise. It comprises multiple features, the most relevant are:

- NAICS Code
- NAICS Description
- Exempt Sales Total
- Taxable Sales Tot
- Gross Sales Total
- Taxable Ratio

The Florida Department of Revenue used industry averages from NAICS to construct its audit lead program. Using the same criteria used by the department to generate the leads, we engineered the Audit Risk dataset with the following features: Taxable Ratio, NAICS Code, Gross Sales Total, NAICS Description, and Gross Income, Tax Paid, Tax Due, and Additional Tax Due. The result of the audit is the class "Audit Risk" which classified the audit risk as High (1) or Low (0).

7. TECHNOLOGICAL STRATEGY

To solve the compliance problem, we utilize computer sciences techniques along with the existing artificial intelligence infrastructures. Since 1955 when U.S. computer scientist John McCarthy coined the term Artificial Intelligence and established the Artificial Intelligence Project in collaboration with Marvin Minsky, AI has been leaping forward, evolving through many approaches, fields, and subcategories, including Logical Reasoning and Problem-Solving Algorithms, Expert Systems, Statistical Inferences and Reasoning, Decision Support Systems, Cognitive Simulations, Natural Language Processing, Machine Learning, Neural Networks, and more. Today, we are delighted that AI has reached the maturity from which we can apply it to solve compliance problems with authority, efficacy, and efficiency. Because AI owes not to be a mystery in the hand of deep-fake bad actors, conspiracy theorists, but an instrument to solve real problem withing communities and societies.

We use Python along with its extensive libraries for data mingling and analysis, and machine learning. Libraries like pandas, NumPy, and matplotlib offer the basis for data manipulation and visualization, while sklearn stands by our side as a one stop shopping center for our machine learning algorithms. Sklearn is such an amazing and deep resource for all types of models, from classifiers to logistic regression, from performance metrics to standard scaler, pipeline to ensemble. We leaned heavily of Sklearn for machine learning, testing, evaluation. For that matter we loaded all the library we could into our project codes.

8. DATA PRE-PROCESSING AND PERFORMANCE METRICS

Indeed, we used modules from sklearn to train our model for classification. We used existing performance metrics, including accuracy, F1 score, precision, and recall, to evaluate our model. After loading the audit risk dataset from memory, we use pandas to transform the csv data file into a data-frame. The data-frame consists of 102920 rows, along with eight features (columns) and one target variable which is the audit risk column.

Data Visualizations: In following best practices and to ensure the dataset suitability from analysis and modeling, we cleaned, preprocessed, and transformed the dataset using pandas' functions, insuring no missing or null values or data anomaly.

To better conceptualize the data, we provide summaries, frequency distribution, and visualization of the characteristics of our data-frame, using pandas descriptive statistics function "describe" and matplotlib. Furthermore, we use the Pearson correlation coefficient to measure the linear relationship between variables.

Feature Selection: To identify key features from the dataset, we first fit features and label (X, y) through the RandomForestRegressor, which is a model from the sklearn ensemble library. Then, we use the feature importances method from pandas' series. From sorting the importance in ascending order, we concluded that "Additional Tax Due" has the highest score: 1.0. So, "Additional Tax Due" is the most important feature in the dataset, which confirms our expectation. To further reinforcing the analysis, we use feature selection from sklearn feature selection. It returns "Additional Tax Due" as the most important feature.

Identifying and selecting the most key features not only simplifies our model, but also it improves performance by reducing overfitting, speeding the training process, and enhance interpretability. The final phase of the training involved pipelining and ensemble, which are computationally demanding. Because feature selection reduces the computational complexity of the model, it truly alleviates the cost of the entire algorithm while increasing predictive accuracy.

Data Transforms: As demonstrated above, while the dataset has eight features, the feature selection utility from sklearn selected one feature "Additional Tax Due" as the most important. We can further simplify the machine learning process while preserving the model's full ability to predict, by reducing the dimensionality of the data. The PCA utility from sklearn decomposition along with the StandardScaler from sklearn preprocessing library come handy for that matter. The resulting normalization and standardization further enhanced the quality and the usefulness of the data by scaling and transforming the data numerical features.

9. TRAINING AND EVALUATION

The systematic and rigorous data preprocessing as elaborated in the sections above left us with a clean, efficient, and enhanced dataset. With all the necessary Python libraries loaded, we proceeded to build, train, and evaluate our model, along the following steps:

1. Partitioning the audit risk dataset into training and evaluation sets, to prevent bias in our model evaluation.

2. Evaluating the performance of the model for each classifier through different metrics, including accuracy, precision, recall, F1-score, AUC, and ROC.

3. Initiating the training of our model using machine learning algorithms for classification, including Naive Bayes, Support Vector Machines, and Neural Networks. The table below shows the results of the initial training.

Model/Metric	**accuracy**	**precision**	**recall**	**F1-score**	**ROC**
Naive Bayes	70.35%	78.80%	94.30%	67.68%	74.87%
Support Vector Machines	81.43%	89.77%	81.43%	100.00%	50.00%
Neural Networks	93.94%	96.26%	96.68%	95.85%	90.72%

As demonstrated in the table above, each algorithm provides a different level of performance. The neural network algorithm offers the highest performance for all metrics, whereas the Naive Bayes algorithm yields fewer satisfying performances. Thus, we completed the initial training without the benefits of feature selection. To verify the advantage of feature selection, we re-trained the model with the injection of feature selection.

The initial training as previously demonstrated in the "Feature Selection" section, Additional Tax Due (score 1.0) is the most important feature, followed by NAICS Code. This confirms our expectation from our knowledge about the dataset. Although NAICS Code has a score of 0.0 in the feature selection, our knowledge of the dataset along with the audit risk algorithm tells us that it is an important distinguishing feature. The NAICS Code identifies industries and sectors. It is indispensable for the model unbiased evaluation. So, we keep it as a feature along with the Additional Tax Due. Then, we re-evaluated our model.

Model/Metric	**accuracy**	**precision**	**recall**	**F1-score**	**ROC**
Naive Bayes	75.09%	82.22%	98.13%	70.76%	82.42%
Support Vector Machines	99.98%	99.99%	100.00%	99.98%	99.99%
Neural Networks	99.48%	99.68%	99.47%	99.89%	98.77%

Indeed, our model evaluations improve for all tested algorithms after reducing the features to just 'NAICS Code', 'Additional Tax Due'. For instance, the accuracy increases from 93.94% to 99.48% for the Neural Networks algorithm, while recall went from 81.43% to 100.00% for Support Vector Machines.

10. PIPELINE, ENSEMBLE, AND CLASSIFICATION

In the final stage of the training, we explore ensemble and pipeline algorithms to boost predictive performance. We implemented three different ensemble-based classifiers, including voting classifier, Stacking Classifier, and Voting Regressor. By combining the predictions from multiple models, ensemble learning algorithms offer better modeling predictive performance.

The different models are themselves fit through pipelines. Pipeline provides built-in functionalities for composite estimators as a sequence of data transformers with an optional final predictor. We implemented three ensemble models each with three different pipelines, for a total of nine pipelines. We then devised a helper function to assist with cross validation scoring.

1. The first ensemble is the voting classifier. It consists of three models, which are Naive Bayes, Support Vector Machines, and Neural Networks

2. The second ensemble is the Stacking Classifier. It combines the following three classifiers: Logistic Regression, Decision Tree Classifier, and K Neighbors Classifier.

3. The last ensemble is the Voting Regressor. Our first two ensembles were classifiers. We figured we could try regression as well as a learning method. We pipeline the three regressors which are the Gradient Boosting Regressor, the Random Forest Regressor, and the Linear Regression, into the Voting Regressor as the ensemble-based algorithm.

After constructing the three pipelines for each of our three classifiers, we use ensemble to combine them altogether for learning, with the goal of making predictions based on majority voting. We injected the 10-fold cross validation to the model (cv = 10) as a hyperparameter to improve the model prediction. This results in better generalization performance than each classifier alone.

The scoring of the voting Regressor is not so promising. The Random Forest Regressor even performs better than the ensemble voting regressor, 98.96% versus 79.56%. The regression is certainly not a good fit for our audit risk dataset. Regression cannot do the job of a classifier. In the other hand, our classifiers perform tremendously well for all computed metrics, first through pipelining them through ensemble 1 and 2, voting classifier and Stacking Classifier, respectively.

CONCLUSION

It is quite elegant to be able to use machine learning algorithms to predict and mitigate sales and use tax audit risk. Our tax audit risk algorithm is a valuable tool for businesses. But also, sales tax professionals, CPA, auditors, and attorneys involved in the auditing process can certainly instrumentalize it to build audit prevention infrastructures as well as defense strategies for their clients.