Building a Predictive Model to Protect Ohio's Citizens

Category: Emerging and Innovative Technologies

State: Ohio

Contact: Katrina Flory, Department of Administrative Services, State of Ohio; Katrina.flory@das.ohio.gov

Project Initiation and End Dates: March 2021 – May 2021

Executive Summary

As a result of the COVID-19 pandemic, there has been a surge in unemployment claims filed by individuals to fraudulently collect unemployment insurance benefits nationwide. In Ohio, the increase in potentially fraudulent claims combined with the unprecedented volume of claims created significant strain on existing business processes within the Ohio Department of Job and Family Services' unemployment insurance program. This then caused a backlog of claims, extending the amount of time it takes for verified claimants to receive their benefits.

In August 2020, the Ohio Department of Job and Family Services (JFS) began work with the Ohio Departments of Health (ODH), Medicaid (ODM), Rehabilitation and Correction (DRC) and the Administrative Services' (DAS) Office of Information Technology (OIT) InnovateOhio Platform (IOP) Data & Analytics Team to address challenges faced in unemployment insurance fraud.

In a significant undertaking, the teams quickly built a scalable framework, including a fully operational Unemployment Insurance Fraud Dashboard, which: flags potential fraud scenarios, allows for continued fraud analysis with drill-down capabilities, and provides direct data access for ad-hoc analysis. Since the deployment of this analytical tool and solution, potential fraudulent monies have been successfully identified for further investigation.

In May 2021, the IOP, JFS Information Services, and JFS Benefit Payment Control (BPC) teams integrated additional functionality related to advanced fraud scoring and fraud detection into the framework and dashboard. Predictive modeling with machine learning capabilities were leveraged to improve the ability to identify and predict high probability of fraud or non-fraud claimants. This cutting-edge technological functionality provides a predicted fraud score and disposition on the dashboard based on statistical probabilities derived from the model.

Machine learning is a strategic and innovative addition to the framework and dashboard; it is designed to improve the accuracy of claim determinations for the BPC team who monitor unemployment fraud claims. As the system continues to identify patterns from the data, its predictive analysis will enhance in accuracy, limiting the considerable level of human intervention previously required to score the likelihood of fraud (or non-fraud). Now, the BPC team will be able to more swiftly move through the adjudication process which determines, identifies, and rejects fraudulent claims and processes valid ones.

Integrating this predictive model into the unemployment insurance fraud framework and dashboard solution is part of JFS' transformation from a reactive unemployment insurance business process and system to a proactive one. By accurately flagging suspicious claims before the money is dispersed, state resources will be safeguarded and more expediently allocated to Ohioans in-need.

Idea

The COVID-19 pandemic brought about unanticipated burdens to existing federal and state business systems and processes, many of which deliver critical services to citizens. Unemployment insurance, a resource providing financial safeguards to eligible individuals by both federal and local governments, was adversely impacted by the pandemic nationwide.

In Ohio, prior to the pandemic, JFS had an established program and process designed to identify potential fraud and disbursement of unemployment insurance monies during normal unemployment trends in the state. However, these processes were not designed to operate in the midst of an unprecedented global health crisis. At the time of the pandemic, Ohio's unemployment rate was at its lowest in a decade. Within a matter of weeks, Ohio's unemployment insurance program was charged with paying millions of Ohioans who were suddenly and unexpectedly out of work. In addition to traditional Unemployment Insurance (OJI) money, JFS also oversaw the administration of new Pandemic Unemployment Assistance (PUA) insurance money as mandated by the federal government's Coronavirus Aid, Relief, and Economic Security Act, or CARES Act.

The dramatic rise in unemployment claims and disbursement of both OJI and PUA insurance monies impacted every part of the claims process within Ohio's unemployment insurance system. As a direct result, the process to determine whether a claim is valid or fraudulent was under significant strain. Determinations were not able to be made accurately or within an average timeframe, leading to fraudulent payments, as well as a mounting backlog of claims, even for legitimate ones.

Total fraud overpayments in the OJI program, from the beginning of the pandemic through the first quarter of 2021, amount to approximately $21 million. Total fraud overpayments in the PUA program from the beginning of the pandemic through February 2021, amount to just under $441 million[1]. These invalid payouts drain fiscal resources, while also inundating state business systems that are designed to protect resources for the unemployed and in-need.

As a result, JFS was presented with a new challenge: how to quickly enhance its current determination process to keep up pace with the new federal regulations and unprecedented demand, while also accurately identifying fraudulent claims. It became critical to leverage advanced technological solutions to pivot JFS' approach and proactively adjust to the severe economic and social burdens, while simultaneously minimizing the level of time required for manual human intervention to effectively preserve state resources for deserving Ohioans.

The initial iteration of the Unemployment Insurance Fraud project included a fully operational dashboard, which flagged potential fraud scenarios on a series of hard-coded (40+) scenarios provided by the Benefit Payment Control (BPC) team. This allowed the team to conduct ad-hoc analysis and easily view the data via a visual dashboard.

---

[1] JFS Numbers as of 5/17/2021

In order to deliver a solution with lasting impacts, JFS and IOP worked together to add additional, cutting-edge enhancements to the framework and dashboard's initial iteration to leverage predictive modeling.

The addition of predictive modeling with machine learning is a key differentiator in this project. The solution now employs advanced statistical analysis to:

- Determine if each of our **current scenarios have a strong (or weak) correlation to actual fraud** (dispositioned as fraudulent claimants)
- **Build a predictive model** that would provide a claimant's likelihood of being fraudulent and continue to adapt/improve based on actual data and statistical analysis on claimant and disposition data
- **Automatically feed new claims through** the predictive model and output results to IOP
- **Provides a predicted fraud score** directly on the dashboard for analysis and BPC team review during the adjudication process, in addition to BPC's calculated fraud score derived from hard coded business scenarios

The predictive model considers actual determined adjudication patterns of claims to predict a claimant of being fraudulent or not fraudulent. The model predicts and scores all claimants on a daily basis to keep information up-to-date and outputs results to the Unemployment Insurance Fraud Dashboard, for real-time analysis. The new predicted score can be used to more accurately identify high-probability fraudulent claimants (or high-probability non-fraudulent claimants).
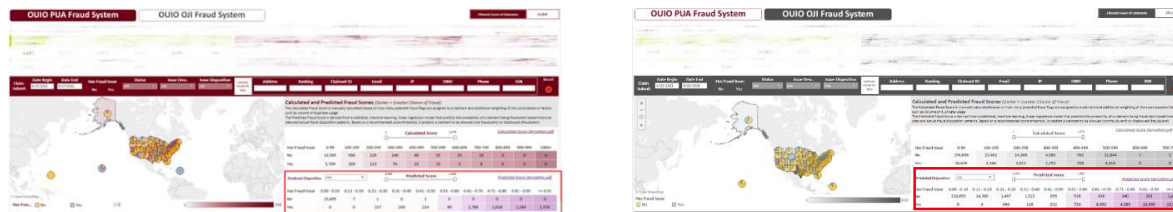

Figure 1: Unemployment Insurance Fraud Dashboard (OJI and PUA) with Predictive Model Score

The addition of this enhancement has the ability to transform JFS' previous claims processes into an advanced analytical and proactive framework to help expediently and confidently validate legitimate claims and reject fraudulent ones, preserving state resources and quickly responding to those in-need.

No state has been immune to the heavy toll the pandemic has had on unemployment resources and insurance programs – in the wake of the pandemic, the United States Labor Department estimates that more than $63 billion dollars[2] has been paid in fraudulent and questionable benefits across the country. States have had to put emergency maintenance safeguards in place to prevent future improper payments to protect their resources.

---

[2] Author: Danielle Serino, "Unemployment Fraud Costs Ohio Millions, Keeps Legitimate Claimants from Getting Their Money," wkyc.com, March 16, 2021, https://www.wkyc.com/article/money/unemployment-fraud-tips-to-avoid/95-893d7b58-0b79-4714-a600-51d031f8f809.

Ohio | Department of Administrative Services

IOP | Innovate Ohio Platform

This project highlights Ohio's response to address the impact of overpayments. However, the nature of this solution is not limited to insurance fraud. By leveraging predictive analytics – the ability to analyze trends based on algorithms that improve automatically based on experience – Ohio may be able to proactively identify trends or solve a variety of problems faced by other state programs.

Implementation

The Department of Administrative Service (DAS) Office of Information Technology's 2020-2022 strategic plan has an acute focus on innovation, in order to improve customer service, reduce cost, and streamline and modernize state systems. The IOP program, within DAS, aligns with this strategy by executing data-driven projects and solutions. State agencies partner with the IOP to leverage its data & analytics platform and suite of advanced tools and capabilities to drive actionable outcomes, such as to aid Ohio's communities; identify inefficiencies, such as fraud; create jobs and predict unexpected events.

JFS and IOP's development of an Unemployment Insurance Fraud framework and dashboard, with predictive analytics to identify fraud and abuse, and to streamline existing processes to improve the lives of Ohioans, is a model example of Ohio's IT enterprise strategy.

The IOP data & analytics team adopted their agile project methodology to rapidly integrate a predictive model with machine learning capabilities into the existing framework and dashboard built during the first iteration of the project. The project development workstream evaluates agency capabilities, optimizes technical solutioning, and implements and monitors success in an agile manner to successfully guide the project from ideation to execution.

The predictive modeling enhancement was implemented in May 2021 and followed a series of project markers to ensure successful integration into the existing framework and dashboard created for the JFS and BPC teams to identify potential fraudulent claims:

- **Analyze, refine, and train** logistic regression model to improve accuracy
- **Run all required claimants** through trusted model
- **Operationalize model** to continuously predict fraud score of new claimants automatically and output to IOP
- **Incorporate** into existing JFS Unemployment Insurance Fraud solutions on IOP (i.e. Tableau dashboards)

This project included several key teams and stakeholders to ensure a successful project. Agency teams included JFS' Delivery Team in Data Management & Analytics and the BPC team, who worked hand-in-hand with IOP's Data & Analytics Delivery Team to land data on the state data platform, develop a logistic regression model based on BPC data, and test and deploy the predictive modeling solution.

This machine-learning feature is still in the early stages of adoption. A current critical strategy to drive adoption includes regular training sprints with end users and the BPC staff, who use the

dashboards as part of their adjudication process. In addition, executive-level demonstrations of the solution's capabilities are currently ongoing with leadership groups from DAS, OIT and JFS to drive awareness of the solution's innate value.

The predictive model was built in two to three weeks and was deployed after four to six weeks of rigorous validation and testing, and then fully integrated into the existing framework and solution. The BPC data derived from the team's adjudication process to identify claimants as fraudulent or non-fraudulent informed the model and algorithms built for the machine-learning feature. This data was already securely shared on the IOP data & analytics platform, which enabled an accelerated delivery of the predictive modeling feature.

The technical architecture for this project is of paramount importance, because the IOP is not single technology, such as a data lake, but a collection of data storage, transformation and analytical tools to adapt to changes in capabilities and design as advances in technology occur. The IOP's data ecosystem and variety of tools allowed for the addition of machine learning capabilities into JFS' Unemployment Insurance Fraud framework.
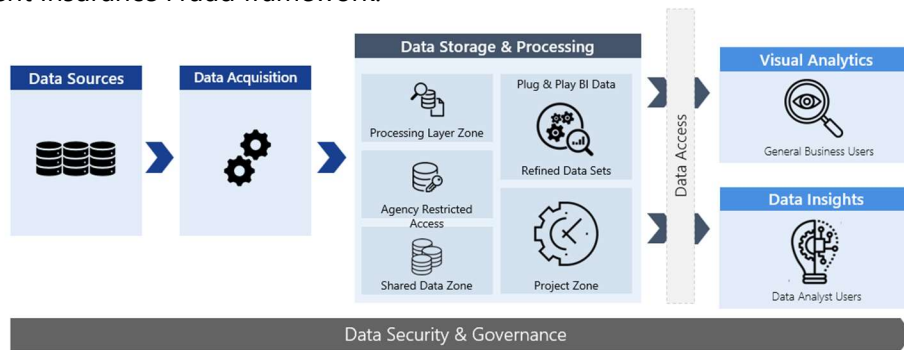
Figure 2: IOP Conceptual Architecture

Alteryx, a code-free data science (predictive, AI, ML modeling) technology that is easy to use and enables data science automation, is fully integrated with IOP technical architecture. This was the key tool used to transform BPC's hard-coded fraud scenarios to a machine learning algorithm to accurately predict fraudulent scores on insurance claims.

Within Alteryx, the following solutions were used to develop the predictive modeling score:
- **Association Analysis** to identify strength of variable relationships to confirmed fraud
- **Logistic Regression** to create statistical algorithm that confidently predicts fraudulent or not fraudulent
- **Score** to assign a probability to a UI claimant on likelihood of fraud based on the logistic regression algorithm

Calculated and Predicted Fraud Scores *(Darker = Greater Chance of Fraud)*
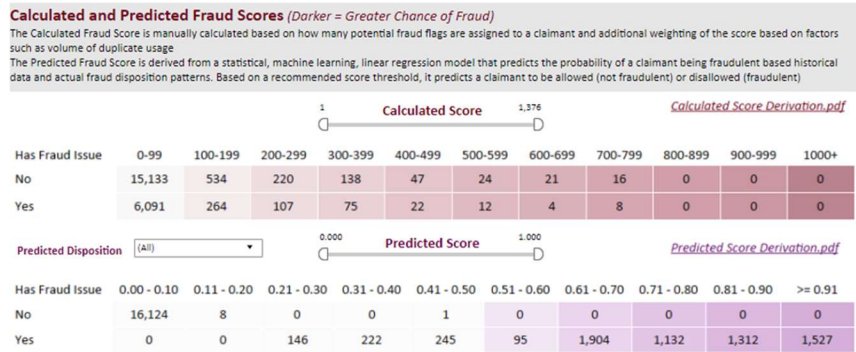The Calculated Fraud Score is manually calculated based on how many potential fraud flags are assigned to a claimant and additional weighting of the score based on factors such as volume of duplicate usage
The Predicted Fraud Score is derived from a statistical, machine learning, linear regression model that predicts the probability of a claimant being fraudulent based historical data and actual fraud disposition patterns. Based on a recommended score threshold, it predicts a claimant to be allowed (not fraudulent) or disallowed (fraudulent)

1   **Calculated Score**   1,376   *Calculated Score Derivation.pdf*

Has Fraud Issue	0-99	100-199	200-299	300-399	400-499	500-599	600-699	700-799	800-899	900-999	1000+
No	15,133	534	220	138	47	24	21	16	0	0	0
Yes	6,091	264	107	75	22	12	4	8	0	0	0

Predicted Disposition   (All) ▼   0.000   **Predicted Score**   1.000   *Predicted Score Derivation.pdf*

Has Fraud Issue	0.00 - 0.10	0.11 - 0.20	0.21 - 0.30	0.31 - 0.40	0.41 - 0.50	0.51 - 0.60	0.61 - 0.70	0.71 - 0.80	0.81 - 0.90	>= 0.91
No	16,124	8	0	0	1	0	0	0	0	0
Yes	0	0	146	222	245	95	1,904	1,132	1,312	1,527

Figure 3: Calculated and Predicted Fraud Scores Visual available in Dashboards

Impact

At the heart of this project, is the fulfillment of JFS' mission to improve the well-being of Ohio's workforce and families, while also ensuring the safety of Ohio's most vulnerable citizens. By employing a predictive model to score fraudulent or non-fraudulent claims, the significant strains that the pandemic introduced on a state system to accurately process and file claims have the potential to dramatically improve. The JFS unemployment insurance process now includes a tool that helps agency staff do the very thing they are passionate about - ensure the safety of Ohio's citizens – delivering untold lasting value and impact.

Now, the backlog of claims is being addressed with the addition of the predictive functionality and fraud score, backed by data. The amount of time it would take to address each of these claims is now anticipated to drastically improve with the help of the predictive model. This is an instant impact to an existing process.

Moving forward, this solution enables a modernized approach to business problems. As the solution matures, the goal is for JFS to continually leverage predictive modeling analytics to allow for a transformation of JFS systems from a reactive to a proactive approach; switching to the new model and prioritizing the utilization of fraud scores to accurately and quickly adjudicate each claim will effectively reserve resources and shorten the claims process.

The solution employed at JFS also has the ability to influence and enable other state resources and systems to leverage machine learning tools and capabilities to predict other trends that may create inefficiencies or abuses, or cause threats to agency resources and processes, creating a positive impact across the State of Ohio. This project highlights the significant and untapped potential of using advanced analytics and emerging technologies to modernize state agencies and existing systems, protecting Ohio's families and communities during unexpected and unprecedented challenges.