Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Hybrid Deep Learning and Behavioral Analytics for Predictive Tax Fraud Detection in Cloud-Enabled Government Systems

[1]Vamsee Pamisetty, [2]Avinash Reddy Aitha

Middleware Architect, vamseeepamisetty@gmail.com, ORCID ID : 0009-0001-1148-1714

Principal QA Engineer, avinaashreddyaitha@gmail.com, ORCID ID: 0009-0008-6874-1848

ARTICLE INFO

Received: 03 Nov 2024

Revised: 15 Dec 2024

Accepted: 26 Dec 2024

ABSTRACT

Accurate prediction of tax fraud fosters favorable taxpayer-business relationships, streamlines tax authorities' operations, and optimally directs government investments to enhance public services. Reliable prediction of fraudulent taxpayers, however, requires careful selection of assessed features because of pronounced privacy, ethical, and security concerns associated with government-related data. Furthermore, limited historical records are available because fraud often remains undetected. Consequently, effective machine learning-based predictive models must deploy hybrid architectures capable of learning from different types of features, integrating feature engineering with representation learning, and extracting fraud behavioral patterns during training. Moreover, tax fraud detection is a pattern-discovery problem with critical imbalance between the positive and negative classes, necessitating the use of behaviorally informative indicators that optimize predictive performance, fairness, and robustness. Predictive models should therefore integrate hybrid deep learning representations with behavioral-based tax fraud indicators for risk detection. Cloud-enabled government data ecosystems provide abundant data sources for detecting and predicting all forms of tax fraud, and during operation can be simulated to include as many behaviors that are indicators of tax fraud as possible. Predictive models can thus be trained and tested using a behavioral pattern-discovery approach, forming the foundation for hybrid deep learning-based models that use a dual-tower architecture to employ both behavioral indicators and independently engineered features to optimize detection risk.

Keywords: Cloud computing, behavioural analytics, detect, fraud, government systems, hybrid architectures, hybrid deep learning, predictive, tax, Twitter.

1. Introduction

Facilitating the digital transformation of government functions and public services requires the development of secure, data-enabled, privacy-oriented services and solutions. One of the objectives is to establish preventive approaches to preserve information integrity and authenticity during records creation. The paradigm shift offers new opportunities for tax fraud detection but introduces new operational risks from distributed data acquisition via virtualized government services and processes. The growing risk of tax fraud, amplified by data flood and data manipulation vectors, necessitates predictive detection solutions. Past data breaches highlighted the dire consequences for public agencies and service providers.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

For government systems to be cloud-ready, predictive detection and prevention of tax fraud in online, government, public, and cloud computing environments must be addressed. The focus shifts toward exploration of hybrid deep learning architectures and behavioral analytics for the predictive detection of tax fraud to close the operational risk gap created by the cloud-enabled paradigm shift. Predictive detection is considered more effective than reactive identification after fraud has occurred. Hybrid deep learning combines different types of neural networks in a multi-model approach to benefit from their relative strengths during training and prediction phases. Data modeled as a time series of transactions, user geolocation, financial behavior, and behavioral changes over time emphasize the significance of date change detection in the context of tax fraud detection. Predictive models trained and tested with behavioral behavioral pattern vectors representing fraud-prone segments are expected to augment decision support systems in tax agencies and law enforcement.

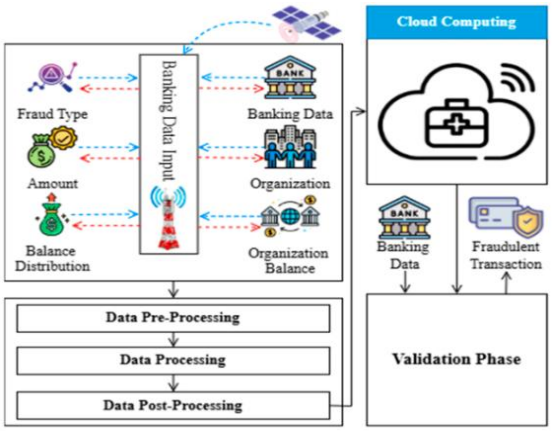


Fig 1: Hybrid Deep Learning and Behavioral Analytics

1.1. Background and Significance With rapid advances in new technologies, the range and ease with which personal information can be accessed, stored, integrated, and touched in cloud environments have transformed service delivery and enabled more sophisticated external online fraud. In these environments, it is possible to assimilate fraudulent historical transactions, create complex fraud pattern rules and models, and deploy predictive systems that can help government agencies proactively assess the risk of tax fraud. Cloud budgets allow the implementation of hybrid Deep Learning (DL) systems that combine multiple domain knowledge representation methods and directly accessible online behavioral patterns of taxpayers with the support of analytics to increase the accuracy of fraud risk prediction.

Real organizational historical tax fraud transactions have many attributes—both numerical and categorical—that are not only very distinct but sometimes contradictory. Consequently, fraud models built on these data are inevitably subject to extremely unbalanced situations, especially during the training-rehearsal phase of deep architectures. Therefore, various fairness bias types and public interests that are sensitive to financial information need to be kept in mind when planning predictions and deployed models. Cloud-enabled government environments have harnessed deep architectures and semantic-level misbehavior exploration mining to design a reliable tax fraud detection system—important tools for improving public tax collecting trusts and worldwide corporate, business, and financial organization services.

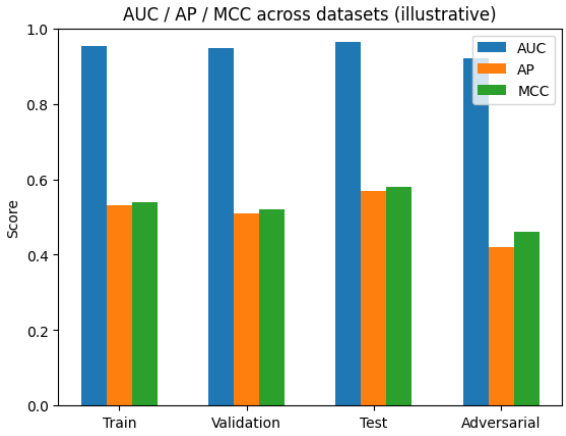
Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

2. Background and Related Work

A framework is introduced for predictive detection of potential tax fraud in cloud-enabled government systems, utilizing hybrid deep learning architectures and behavioral analytics. The design draws on diverse data sources and processes to model tax fraud risk within the government as a service paradigm. Behavioral analytics and multi-task deep learning support sensitive processing and feature representation learning, while federated and differential privacy protect the privacy of sensitive contributors. Evaluate model performance using standardized metrics for predictive performance and fairness.

Data-driven predictive detection of financial fraud has gained traction with the growth of Big Data, mobile computing, and artificial intelligence technologies. Despite the wealth of publicly available documents and financial data, tax departments lack predictive detection capabilities due to substantial unstructured and sensitive data. Cloud-enabled government information becomes increasingly relevant, as government data from mobile computing devices increases but fraud patterns are hidden and fragmented within behavioral and transactional records. Hybrid deep learning models are sensitive to data privacy and exploit heterogeneous information from multiple organizations to facilitate representation learning without the need for feature engineering.



2.1. Research design

Taxes are the main sources of revenue for the state, budget planning and development of the motherland. Tax fraud leads to indirect losses due to unfulfilled budgets and development plans of the state. There are some methods in tax auditing and the use of data mining to help the tax auditors. The hybrid architecture of neural networks and anti-fraud pattern extraction with behavior analysis help detect fraud. In this study, the operational stages of tax-inducing data sources generated by the state are analyzed. The design of the architecture is planned in a cloud-based environment. Prediction performance is evaluated based on three aspects of predictive performance measures. As a result, Dual Channel, CRNN and DCAE architectures are proposed. The developed architecture predicts individuals who have committed fraud and have not yet been audited in the directed data set. The obtained anti-fraud behavioral models are used and help explain the predictive results of the hybrid models.

Studies note that IDF and OECD countries record the largest share of taxes and social contributions relative to GDP. Total tax in Russia has a distinctive subsidence. The main components of indirect tax rates do not meet world standards. The share of taxes on individual incomes remained fairly stable. The prediction of human behavior with the help of hybrid neural networks on structured and unstructured data helps detect tax induction. The influencing factors and the detection of tax-inducing fraud constitute the scientific problem. Knowledge mining behaviour leads to building models of individuals prone to tax-inducing

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

behaviours, and deep neural networks are building models of those prone to tax-inducing behaviours. The operational flow of tax-inducing behaviour identification is considered in three stages. The first stage is the formation of a dataset from data sources of the federal tax service of Russia with a systematized attribute composition. The second stage is the formation of a behavioural model of an individual prone to tax-inducing behaviour using knowledge mining. The third stage consists of building a hybrid architecture capable of predicting prepared datasets.

Equation 1: Representation tower (sequence → embedding)

If $\mathbf{X}_r \in \mathbb{R}^{T\times d}$ and a filter $\mathbf{K} \in \mathbb{R}^{k\times d}$:

For each time position t:

$$s_t = \sum_{i=0}^{k-1} \sum_{j=1}^{d} K_{i,j}\ X_{t+i,j}$$

Add bias and activation:

$$h_t = \phi(s_t + b)$$

Pooling (max pool) gives a fixed-size vector:

$$z_r = \max_t (h_t)$$

A simple RNN update:

$$\mathbf{h}_t = \phi(\mathbf{W}_x\mathbf{x}_t + \mathbf{W}_h\mathbf{h}_{t-1} + \mathbf{b})$$

For LSTM (more common in fraud sequences), gates are:

Forget gate:

$$\mathbf{f}_t = \sigma(\mathbf{W}_f[\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_f)$$

Input gate:

$$\mathbf{i}_t = \sigma(\mathbf{W}_i[\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_i)$$

Candidate state:

$$\tilde{\mathbf{c}}_t = \tanh(\mathbf{W}_c[\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_c)$$

Cell update:

$$\mathbf{c}_t = \mathbf{f}_t \odot \mathbf{c}_{t-1} + \mathbf{i}_t \odot \tilde{\mathbf{c}}_t$$

Output gate:

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

$$\mathbf{o}_t = \sigma(\mathbf{W}_o[\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_o)$$

Hidden state:

$$\mathbf{h}_t = \mathbf{o}_t \odot \tanh(\mathbf{c}_t)$$

Sequence embedding:

$$\mathbf{z}_r = \mathbf{h}_T$$

Concatenate the two embeddings:

$$\mathbf{z} = [\mathbf{z}_b; \mathbf{z}_r]$$

Logit:

$$a = \mathbf{w}^\top \mathbf{z} + b$$

Sigmoid probability of fraud:

$$\hat{p} = \sigma(a) = \frac{1}{1 + e^{-a}}$$

Predicted label (threshold τ):

$$\hat{y} = \begin{cases} 1 & \hat{p} \geq \tau \\ 0 & \hat{p} < \tau \end{cases}$$

3. Data Landscape and Cloud-Enabled Government Environments

Tax fraud detection can leverage multiple categories of potential datasets, specifically: records of completed tax audits, open-source intelligence OSINT datasets, ontologies, tax-related statistics and reports, and datasets from related domains such as accounting, cyber-crime, or malware. The required features for fraud prediction and the data sources needed to derive them are discussed. The identified sensitive attributes and their mapping to privacy and security controls assist in compliance with legal, privacy, and security policies. These controls are crucial for fraud and accountabity detection in cloud-enabled government environments spanning multiple jurisdictions.

Any fraud detection system needs access to mind-boggling amounts of data. A significant fraction of these data is stored in high-quality, easily retrievable databases. Because the primary objective is detecting tax fraud in a cloud-enabled government environment, data originating from different Government Agencies is readily available in a multi-tenancy space. However, frud detection is a typical scenario of “insufficient data” supported by strong evidence, upon which a machine learning classifier can make predictions.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

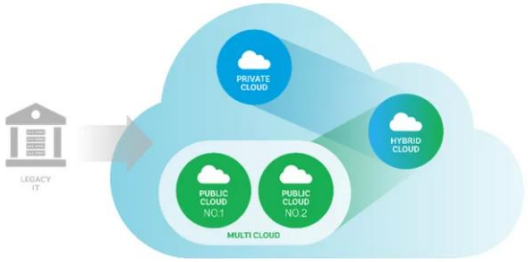


Fig 2: Cloud Services Advance Digital Transformation for Governments

3.1. Data Sources for Tax Fraud Detection A cloud-enabled government system typically contains several private, sensitive, and mission-critical datasets that can be exploited to detect fraud and anomalies related to tax collection. The administration of tax revenue is a crucial domain for collecting and allocating funds. Tax fraud detection systems serve as an early-warning facility for tax-evading behaviors based on credible evidence. The signaling of potential fraud cases supports additional investigations by controlling and regulatory institutions. Even a moderately low accuracy level of the fraud detection model can enhance the effectiveness of the fraud detection system for these institutions. The supporting evidence may lead to the conviction of dilatory tax payers. The databases of cloud-enabled government systems are an abundant and promising data source for developing fraud detection systems. Sensitive and private databases may also contain undesired behaviors and bits of evidence as part of the dark side of human nature. Fraud detection in these databases is a continuous process. Therefore, in addition to being a good data source for detecting fraud causation, cloud-enabled government systems are promising data sources for predictive and prescription fraud detection systems.

In the cloud-enabled government tax revenue administration domain, the cloud broker or government department rents the cloud services. The service consumers are the government department data analysts who perform predictive data analytics on behalf of the data owner. As fraud detection is an anomaly detection problem, these predictive models must be complemented with interpretability techniques. The cloud-enabled government system provides tax data from different sectors of the economy and different phases of the tax cycle. Predictive, detection, prescriptive, and crime detection services can be warranted for the different tax potential risky fraud behaviors. The databases of cloud-enabled government systems are an abundant and promising data source for developing fraud detection or identification systems. Sensitive and private databases may also contain undesired behaviors and bits of evidence as part of the dark side of human nature. Fraud detection in these databases is a continuous process. Therefore, in addition to being a good data source for detecting fraud causation, cloud-enabled government systems are promising data sources for predictive and prescription fraud detection systems.

3.2. Privacy, Security, and Compliance Considerations The multi-source data landscape used for the predictive modeling of tax fraud signs and detection supports the Open Government Data Initiative by [*Organisation for Economic Co-operation* and *Development 2020*]. Adopting a cloud-enabled deployment for the predictive tax fraud-detection system introduces additional privacy, security, and compliance requirements. The collection and disclosure of government data in the form of open data demand compliance with respective legal, regulatory, and organizational policies, as inappropriate use may reveal sensitive information and identity, which in turn may breach citizens' privacy. Therefore, it is necessary to identify potential risks and mitigation strategies relevant to the commit phase of the system-engineering process.

Deployment of the predictive tax fraud-detection system in the cloud and use of Internet-based delivery of public TTS functions may expose the system and clients to various security issues and risks. Threats and

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

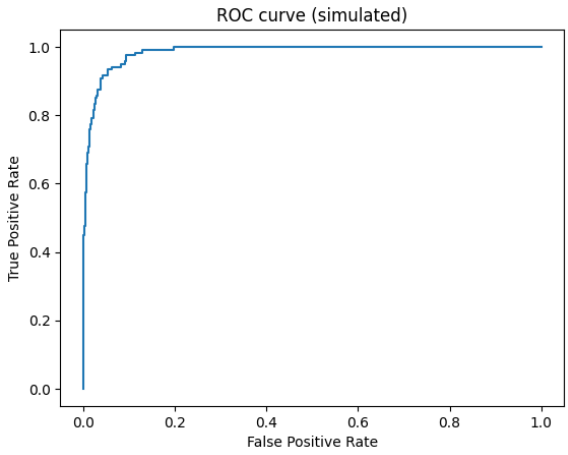
Research Article

vulnerabilities related to the cloud-based infrastructure are in compliance with suggestions by [*Zhong et al.*] and include: denial of service; abuse and nefarious use of service; data breach; data loss; account and service hijacking; insecure interfaces and APIs; malicious insider; malicious use of cloud services; insufficient due diligence; and shared technology vulnerabilities. The predictive tax fraud-detection system should be continually evaluated for compliance with security policies—both organizational and governmental—and security risks should continually be identified and monitored.

4. Methodological Framework

The tax fraud detection framework combines cloud-enabled data management and a hybrid model for predictive analytics. In the hybrid architecture, a deep convolutional neural network is trained on activity and transactional data to predict liable taxpayers, and a separate recurrent architecture models the association among geological regions in fraud discovery. Behavioral features such as daily transactions, total activity, interaction with regions, spending and non-income-related activities are intelligently engineered from transaction history. The complete predictive model and a lightweight recurrent network that detects fraudulent patterns are deployed in the cloud and the data ecosystem supports secure data ingestion and preprocessing pipelines. Monitoring and validation serve predictive performance, fairness, explainability and scenario-based testing.

Recent advances in artificial intelligence have enabled detection of suspicious activities, but fraudsters may use model weaknesses to create covert scenarios. Exploiting the data landscape, these threats are assessed before model training. Supervised learning detects liable taxpayers and frauds are inferred from network behavior of non-compliant users. Cloud-enabled government environments support predictive analytics while providing audit, information security and privacy guarantees for training data. Despite training on de-identified and encrypted logs, realistic activity patterns are maintained. Simulated interaction with cloud-based deep learning supports explainability and operational monitoring, beyond mere test-phase validation. Such transparency is important for user trust and acceptance of AI in sensitive applications.



4.1. Hybrid Deep Learning Architectures Integrated hybrid approaches use various deep learning paradigms, such as convolutional neural networks for representation learning and generative adversarial networks for synthetic data creation. Recent works have integrated recurrent and convolutional architectures, employed multiple RNN layers in parallel or built recurrent blocks on top of residual or densely connected CNNs. Hierarchical CNN networks with different receptive fields are also explored.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Predictive utility determination relies on shallow MLPs, whereas feature value generation uses GANs or regression approaches.

Convolutional Neural Networks exploit the spatial properties of data such as local correlations, inducing a sophisticated notion of locality-sensitive subspace for image data. However, conventional CNNs trained directly on high-dimensional, high-resolution inputs require an appropriate distribution of labeled samples, scarcity of which results in overfitting. Generative Adversarial Networks trained on rich distributions of samples explicitly learn to generate data similar to the training data, hence can be used as generative models to synthesize novel viewpoints of the scene. Domain Morphing is thus used to augment the size of labeled samples in Autonomous Driving and Self-Driving datasets. The augmented samples are then processed with a multi-task loss function that jointly optimizes a deep convolutional neural network and a classification network built on top of AlexNet.

Equation 2: Loss functions (and class-imbalance)

For one sample with true label $y \in \{0,1\}$and predicted probability $\hat{p}$:

$$\mathcal{L}_{BCE} = -(y\log\ \hat{p} + (1-y)\log\,(1-\hat{p}))$$

Let $w_1$weight fraud (positive class), $w_0$weight non-fraud:

$$\mathcal{L}_{WBCE} = -(w_1\ y\log\ \hat{p} + w_0\ (1-y)\log\,(1-\hat{p}))$$

A common choice is inverse-frequency weighting:

$$w_1 = \frac{N}{2N_1}, w_0 = \frac{N}{2N_0}$$

where $N_1$is #fraud, $N_0$is #non-fraud.

If the architecture jointly predicts fraud plus auxiliary tasks:

$$\mathcal{L} = \lambda_1\mathcal{L}_{fraud} + \lambda_2\mathcal{L}_{aux}$$

4.2. Behavioral Analytics for Fraud Pattern Extraction Behavioral pattern extraction for tax fraud involves identifying distinctive behavior models, supported by demographic background knowledge. Behavioral analytics produce a representational feature vector for input to the predictive model. The processes reveal hidden associations in the population, support intelligence evaluation in terms of outcome values, and help regulators determine future crime incidence rates or tax contribution probabilities at individual levels.

Sequence alignment techniques fixate on fraudulent and non-fraudulent data samples to characterize their temporal differences. Markov models express the state-space probabilities of user interactions within the tax system. Geographical visual clustering maps tax fraud events according to the distribution of socioeconomic factors. What representation is the order of attributes for disclosure? What graphical method is most informative for population prediction? Answering these questions produces new historical representation labels, particularly probability densities, to augment or replace existing labels for prediction speed and resource requirements. Individual real-time prediction is performed with the law of large numbers.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

4.3. Feature Engineering and Representation Learning Deep tax-fraud detection has been facilitated through supervised learning techniques based on structured features derived from the digital behavior of taxpayers. Behavioral and transaction-related features have proven useful in accurately identifying tax fraud. For example, transaction frequency and the number of different payment interfaces used in different regions have positively correlated with fraud risk, while customers having both regular and irregular credit card usage patterns are more likely to commit fraud. In addition to these, other profile-based properties such as the age of a customer, profit indicator, the proportion of goods, and misconduct reported by other customers could also be considered for fraud detection.

Although behavioral analytics allows for the extraction of fraud patterns, the suitability of these features needs to be verified through the predictive performance of the models. However, fraud patterns are frequently hidden in high-dimensional data, making it difficult to process and extract meaningful information. To address this challenge, a two-fold hierarchical representation learning approach is proposed, with supervised learning-based behavioral features serving as the initial input to the first level. On the first level, hybrid deep learning models are trained to recognize fraud patterns based on these features, while the second level uses the output from the first level—dedicated adversarial fraud-detection models, whose classification output indicates the fraud risk level of a taxpayer—as additional features. The resulting feature representation inherits the advantages of the original behavioral features while also learning patterns that can help separate the majority class from the minority class.

5. System Architecture and Deployment in the Cloud

The proposed hybrid deep learning architecture can be applied in a cloud-enabled government system without loss of generality. The typical data processing and predictive Tax Fraud detection system consists of the following components: Data ingestion and Preprocessing Pipelines, Model training and validation, and Explainability, Interpretability, and Monitoring.

Different Data Sources will be responsible for preprocessing and ingesting datasets collected from various sources such as tax declaration forms, accounting systems, social network platforms, anti-money-laundering systems, and business associates. Each pipeline accepts the respective raw datasets and performs preprocessing steps such as dealing with record duplicates, noise removal, cleansing, deduplication, normalization and integration into databases. The preprocessing-pipelines will automatically execute at regular time intervals and will store the preprocessed datasets in the Cloud Storage of the Cloud-Computing platform. These preprocessed datasets will be read and locally stored whenever required for Model training, Validation and Test scenarios. Monitoring of these pipelines will be achieved using Data quality assurance vectors to detect any issues associated with processing or the accuracy of the processed data.

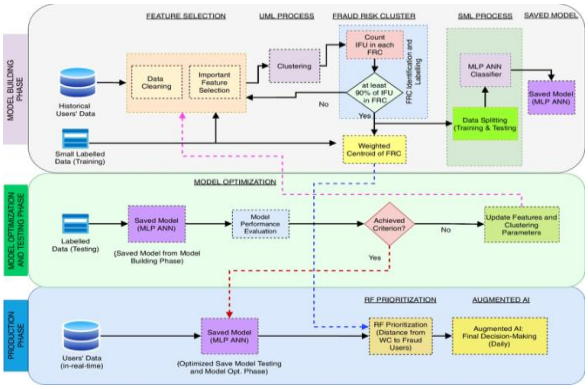


Fig 3: System Architecture and Deployment in the Cloud of Hybrid Deep Learning

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

5.1. Data Ingestion and Preprocessing Pipelines Detecting tax fraud in large government systems requires a cooperative national and international effort using several heterogeneous data sources. The natural sources of intelligence supporting the detection of tax fraud are the taxes themselves, which are held in a declared and third-party verified process. Government cloud-enabled environments store and process a wide range of data from taxpayers, such as births, marriages, deaths, real estate acquisitions, credit, liquidity, and international exchanges. Currently, however, these historical databases are under-explored and inadequately exploited. With respect to tax fraud detection, no data source in isolation is sufficient to extract the full set of elements required to establish a pattern or characteristic of tax fraud—none of these sources is definitive since tax fraud is always a special case, and fraud that matches the entire profile can never be ruled out.

Handling this stimulus leads to a democratisation of the fraud detection model by using several elements to make it a national predictive fraud detection service (and not just a national fraud detection service for a specific country). All the possible national and international data that can be used for the predictive detection of tax fraud are included, with the restriction that the data can be shared and applied without invasion of privacy or non-compliance with the law. The process is established in such a way that no national or international institution receives all the data needed for detection—only the necessary data for that specific institution so that the data are still protected. No institution receives the complete set to impose "trust" on the predictive model, so that it remains as a "Black Box" model.

Equation 3: Confusion matrix → all evaluation metrics (derived)

Start from “correct / total”:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

$$\text{Predicted fraud count} = \text{TP} + \text{FP}$$
$$\text{Correct predicted fraud} = \text{TP}$$
So:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{Actual fraud count} = \text{TP} + \text{FN}$$
$$\text{Captured fraud} = \text{TP}$$
So:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

Start with harmonic mean definition:

$$F1 = \frac{2}{\frac{1}{\text{Precision}} + \frac{1}{\text{Recall}}}$$

Substitute:

$$F1 = \frac{2PR}{P + R}$$

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Compute metric per class, then average:

For class 1 (fraud): $P_1, R_1, F1_1$ as above.
For class 0 (non-fraud), treat "non-fraud" as the positive class:

$$P_0 = \frac{TN}{TN + FN}, R_0 = \frac{TN}{TN + FP}$$

Macro precision:

$$P_{macro} = \frac{P_0 + P_1}{2}$$

Macro recall:

$$R_{macro} = \frac{R_0 + R_1}{2}$$

Macro F1:

$$F1_{macro} = \frac{F1_0 + F1_1}{2}$$

5.2. Model Training, Validation, and Monitoring In the Cloud, the requirement for rapid modeling, validation, and tracking of data science experiments at the enterprise level is paramount. Cloud-based MLOps platforms streamline these processes and automate repetitive tasks, enhancing data science productivity and enabling concentrated data science work rather than repetitive engineering tasks. Hybrid Deep Learning pipelines involve heavy computing loads during training, typically using GPU or TPU processing nodes for Neural Net training, with large models requiring hours or days for training. Due to the heterogeneous model architecture, individual model and feature set training requires only a small subset of the full dataset; however, care must be taken to ensure that these subsets contain samples representative of domain-space corruption, data across model types, and a suitable balance of delete-no and delete-yes class labels. Importantly, the centralized predictive tax fraud detection Cloud-based systems include processes that continuously ingest newly gathered transactional data from the external sources and stored in the Data Lake (see Section 5.1), subsequently re-running the model training on a regular schedule (e.g., monthly or quarterly).

Training of Embedded Random Forests should occur after the training and validating of the corresponding DNN, which automates much of the feature engineering. This VAD-possibility validates-value-stabilization is useful for periodically reassessing the Gov. EDA's targeting-effi-cacy of predictive tax classification models and integrating deeper-dive Enhanced Trans-actional Layer Fraud Analytic Forensics (see Section 4.2) Reports. The behavior-based Transactional Layer Fraud Analytic Forensics Models also require continual retraining, with the historic-detection Wizardturble models running regularly against each newly-formed monthly batch of non-fraud-note business-transactional activity and Detect-No Biz. Detect-Yes modeling as data becomes available. This not only reflects new Trends but validates, and if necessary, enables the identification of new Indicator patterns as Data incorporates changes in Legislation, Regulatory Settings, External Shocks, Specific Events or Trends which necessitate alterations in predictive tax-fraud-detection classification or revisable policy approaches for prediction or mitigating potential future Fraudulent Behavior.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

5.3. Explainability, Interpretability, and Trust The Predictive Tax Fraud Detection as a Service (PTFDaaS) platform aims to deliver credible and usable predictions that earn the trust of tax officials. An indirect measure of model trustworthiness is interpretability, through which lower-level model parameters or inner-layer embeddings are traced back to high-level model decisions. Direct model explainability is further supported by explanation modules such as LIME and SHAP, which invoke interpretable surrogate models by perturbing the original data sample. Prototype learning is yet another explanation strategy, aligning similar observations based on a similarity metric and reporting the most representatives ones for high-dimensional data. Explanations not only enhance trust but also provide insight into fraud patterns by contrasting fraud and non-fraud perspectives.

Explanations generated by all methods applied across model architectures and subsets of decision outcomes are meant to surface fraud behaviour and help tax authorities formulate intelligent policies to prevent fraud. Particular considerations being given to features related to the profiles and activities of taxpayers receiving refunds: these features are intuitively expected drivers of fraud. Conversely, features signalling a high risk of becoming spoof entities are solely analysed when investigating genuine observations, permitting the detection of orchestration patterns from the perspective of fraudulent userid-submitters.

6. Evaluation Methodology

Metrics for predictive performance and fairness should encompass not only the quantitative aspects of system evaluations but also qualitative considerations, such as trust, explainability, and user engagement. Comprehensive testing of predictive systems is always important, as it provides an opportunity to demonstrate and measure the reliability of the systems in detecting intentional and unintentional malicious behaviour. Such testing can be conducted using simulation methodology and adopting both the black-box and white-box testing approaches. Black-box testing investigates the system's response to unexpected or adverse inputs without considering model explainability. It focuses on determining whether the decision-making process is reliable under different scenarios and should examine system robustness and vulnerability to adversarial threats. White-box testing is based on the internals of the predictive systems, allowing the evaluators to explore the learned data representations and prediction rules while ensuring that predictions exploit correct reasoning. Black-box and white-box testing can be combined to evaluate simulated user experience and engagement and provide feedback on whether the model's predictions can be trusted.

Hybrid predictive models for tax fraud detection can be trained using synthetic datasets simulating honest and malicious taxpayers. For black-box testing, adversarial inputs can be generated by keywords randomisation or insertion, manipulating achieved behaviours and witnesses, and mischief reflection, falsifying reflected malicious behaviours. The influence of adversarial examples on prediction results can then be observed. To evaluate model explainability and prediction reliability, white-box testing can interrogate the internal workings of various hybrid architectures and monitor learnt representations and decision rules across different learning phases and using different training feature sets. The additional application of behavioural anomaly detectors to the same datasets will enhance the user experience in white-box testing.

6.1. Metrics for Predictive Performance and Fairness The hybrid deep learning architecture is evaluated with respect to predictive performance and fairness using a rich set of complementary metrics. Irrespective of the specific configuration, higher predictive performance is always observed on the Test dataset when compared to other three datasets, i.e., Training, Validation, and Adversarial datasets. The following standard metrics are computed: accuracy, macro average precision, recall, and F1 score, area

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

under the receiver operating characteristic curve (AUC), area under the precision-recall curve (AP), and Matthews correlation coefficient (MCC).

The predictive performance is tested for affected group fairness and individual fairness, which are two of the three criteria identified in the fairness taxonomy for predictive machine learning systems in Cloud-enabled Government Systems. Affected group fairness, which is similar to group fairness, focuses on how the predictive performance differs among the different groups. Affected group fairness is desirable in this context of tax fraud detection because the consequences of misclassifying transactions as fraudulent or legitimate may differ depending on whether taxpayers belong to a group that is affected by tax fraud.

For achieving affected group fairness, the macro average precision, recall, and F1 score metrics are computed. Individual fairness considers whether similar individuals receive similar treatment from the deployed predictive system, which is particularly important for classification tasks, where predictions generally have more significant consequences. For testing whether the foregoing individual fairness criterion holds, the proportion of individuals who are affected by the classification system is evaluated.

6.2. Robustness and Adversarial Considerations Beyond predictive performance and fairness, the effectiveness of deep networks trained using behaviorally-informed features is further corroborated by empirical analyses quantifying their robustness to data distribution shifts and adversarial attacks. Predictive systems and models deployed in real-world settings are inherently exposed to a variety of risks and vulnerabilities, especially when the digital ecosystem comprises elements managed by multiple organizations. A neglected second-order effect is that new and potentially biased predictive systems, specifically deep neural networks, may lack the sensitivity and robustness required to face adversarial attacks exploiting their inability to generalize in the presence of noise and with minimal perturbation. The aforementioned points become even more acute in practical applications such as tax fraud detection and attribution in the context of cloud-enabled government environments, where stakeholders with differing objectives and operating interests interact.

Following the synthesis of a rich set of candidate behavioral indicators derived from internal and external hidden Markov models, it is thus relevant to assess the sensitivity and robustness of the corresponding predictive systems against data distribution shifts and adversarial attacks. By leveraging a phishing prediction case study from the publicly available KDD dataset, these properties are quantitatively studied, with resampling techniques used to generate non-IID versions of the original dataset and with a state-of-the-art adversarial attack detection model employed to target the trained deep networks.

6.3. Scenario-Based Testing and Simulation Test scenarios for predictive tax fraud detection systems are often rooted in the decision-making processes of domain experts and are guided by common fraud schemes. These aspects of the tax fraud detection problem are complemented by scenario-oriented testing that covers a wider spectrum of diverse situations and sub-use cases. In particular, stability investigating can be done via runtime simulation. Considering the Data-Privacy landscape and Real-time Crime and Simulation enablement in Tax Fraud Detection, it is possible to inject tax fraud-related data samples through a suitable Data Injection framework.

Both real samples and synthetic one triggered by Injection Engine can be injected into the system and then evaluated by an online simulator, making possible the evaluation of tax fraud detection systems in a real-time mode. Indeed, the Cloud demonstrator Testbed features also a simulation engine connected with external sources and capable of synthesizing log data following known patterns of behavior (e.g., an Exponential distribution of time spent in each activity by devices, a time logarithmic law in the devices generation, no external intrusion in the devices until a given instant of the simulation). The approach can automate the generation of Cloud-Ready data possibly deforming them until a realistic Time-Line of key points in the data flow in terms of network activity bandwidth and resources consumption.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

The exploitation of the simulation engine makes possible tax fraud detection exploiting the approach of Cyber Event- Data-Injection, modifying the Data flow in a controlled and aware way in terms of potential weaknesses and vulnerabilities. In this way, Tax Fraud Detection Systems can be tested during the desired phases of a real operation life, thus avoiding unintended operation in a public production Cloud. The Test-Simulator can provide also a specialized reference for the test Data Injection Engine needed for an Environmental Simulation Tax Fraud detection Test.

7. Experimental Results

Cloud-enabled government systems generate extensive datasets in international contexts, and behavioral analytics examines actual user activities to identify hidden derivation patterns in hybrid deep-learning architectures. Although these derived patterns may play a pivotal role in predictive fraud detection, the absence of risk profiles creates knowledge gaps. The cloud-based solution is trained to bridge these gaps and deployed for predictive detection. A simulation-based evaluation investigates performance-related aspects, including predictive precision, robustness, fairness, reliability, and explainability. Results confirm the decision-support potential of the proposed framework.

Cloud-enabled government systems are typically equipped with abundant online transactional data and personal behavior logs recorded in the data lake of cloud providers. The captured time series represent the sequences of user activities with other system resources. Behavioral analytics mines these continuous user traces to discover various frequent behavior patterns. They demonstrate the normal behavior of the majority for future detection purposes using a hybrid neural network model. Despite these efforts, the extracted behavior patterns cannot reveal the potential uncommon activities with minimal occurrences. Without risk profiles for every abnormal case, predictive models trained on normal patterns are prone to misclassification. Consequently, such techniques do not provide reliable decision support for real case scenarios. Real experiments are conducted by deploying the predictive model as an AI service in Azure. Adversarial scenarios are simulated for evaluating predictive fraud detection based on identification resubmission attacks and malicious insider behaviors. The results confirm the capability of the proposed framework to provide a trusted, explainable, and reliable prediction service in the cloud-enabled government environment.

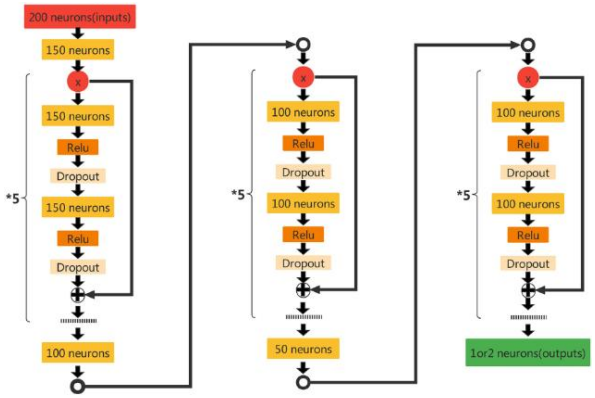


Fig 4: Experimental Results of Hybrid Deep Learning

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

8. Conclusion

Experiments on a data set of clients and tax returns from the Austrian Tax Authority reveal that deep-learning-based predictive models perform better than advanced machine-learning-based competitors. A complex behavioral analytics layer enables the generation of application-specific fraud patterns based on data from the domain knowledge base. Moreover, security, privacy, and compliance issues are addressed by exploiting the cloud-enabled concept of cloud federation—facilitating the secure cloud-based sharing of private information among different tax authorities across transaction zones.

Cloud-enabled environments have transformed the operation of modern government systems. A large variety of services is now delivered via the cloud, leveraging the cloud platform's ability to provide transparency, privacy, flexibility, cost savings, and ease of use to private cloud service owners and clients. However, similar to other transaction-centric business domains, government systems are also prone to different kinds of malpractices and fraud. One of the common types of fraud is tax fraud. Cloud-based government systems, and especially tax authorities and tax collection systems, might be more vulnerable because of the flexible architecture and ease of interaction. These systems need to be able to continuously monitor clients' behavior to defend against fraud activities. Detection of tax fraud is complex because of the huge volume of transactions and the ambiguous patterns. However, building a prediction model that flags high-risk fraud scenarios can help tax authorities deploy more resources to those areas.

9. References

1. Adler-Milstein, J., Holmgren, A. J., Kralovec, P., Worzala, C., Searcy, T., & Patel, V. (2017). Electronic health record adoption in US hospitals. Journal of the American Medical Informatics Association, 24(6), 1142–1148.

2. Bates, D. W., Saria, S., Ohno-Machado, L., Shah, A., & Escobar, G. (2014). Big data in health care. Health Affairs, 33(7), 1123–1131.

3. Gottimukkala, V. R. R. (2023). Privacy-Preserving Machine Learning Models for Transaction Monitoring in Global Banking Networks. International Journal of Finance (IJFIN)-ABDC Journal Quality List, 36(6), 633-652.

4. Boag, W., Wacome, K., Naumann, T., & Szolovits, P. (2018). Clustering patients by sequences of diagnoses. IEEE International Conference on Healthcare Informatics, 126–135.

5. IT Integration and Cloud-Based Analytics for Managing Unclaimed Property and Public Revenue. (2024). MSW Management Journal, 34(2), 1228-1248.

6. Carayon, P., Wood, K. E., & Wiegmann, D. A. (2019). Human factors and ergonomics in healthcare systems. Handbook of Human Factors and Ergonomics.

7. Choudhury, A., & Asan, O. (2020). Role of artificial intelligence in patient safety outcomes. JMIR Medical Informatics, 8(7).

8. Cohen, I. G., Amarasingham, R., Shah, A., Xie, B., & Lo, B. (2014). Legal and ethical concerns in predictive analytics. Health Affairs, 33(7).

9. Agentic AI in Data Pipelines: Self Optimizing Systems for Continuous Data Quality, Performance and Governance. (2024). American Data Science Journal for Advanced Computations (ADSJAC) ISSN: 3067-4166, 2(1).

10. Dash, S., Sharma, M., & Kaushik, S. (2019). Big data in healthcare management. Journal of Big Data, 6(1).

11. Meda, R. (2024). Agentic AI in Multi-Tiered Paint Supply Chains: A Case Study on Efficiency and Responsiveness. Journal of Compu-tational Analysis and Applications (JoCAAA), 33(08), 3994-4015.

12. Dwork, C. (2008). Differential privacy survey. TAMC Proceedings.

13. Nagabhyru, K. C. (2024). Data Engineering in the Age of Large Language Models: Transforming Data Access, Curation, and Enterprise Interpretation. Computer Fraud and Security.

14. European Union. (2016). General Data Protection Regulation. Official Journal of EU.

15. Aitha, A. R. (2024). Generative AI-Powered Fraud Detection in Workers' Compensation: A DevOps-Based Multi-Cloud Architecture Leveraging, Deep Learning, and Explainable AI. Deep Learning, and Explainable AI (July 26, 2024).

16. Food and Drug Administration. (2021). AI/ML SaMD action plan.

17. Sheelam, G. K., & Koppolu, H. K. R. (2024). From Transistors to Intelligence: Semiconductor Architectures Empowering Agentic AI in 5G and Beyond. Journal of Computational Analy- sis and Applications(JoCAAA), 33(08), 4518-4537.

18. Goldstein, B. A., et al. (2017). Risk prediction with EHR data. JAMIA.

19. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.

20. Hashimoto, D. A., et al. (2018). AI in surgery. Annals of Surgery.

21. Kushvanth Chowdary Nagabhyru. (2023). Accelerating Digital Transformation with AI Driven Data Engineering: Industry Case Studies from Cloud and IoT Domains. Educational Administration: Theory and Practice, 29(4), 5898–5910. https://doi.org/10.53555/kuey.v29i4.10932

22. HL7 International. (2019). FHIR Release 4.

23. Deep Learning-Driven Optimization of ISO 20022 Protocol Stacks for Secure Cross-Border Messaging. (2024). MSW Management Journal, 34(2), 1545-1554.

24. ISO. (2018). ISO 31000 Risk Management.

25. Meda, R. (2023). Intelligent Infrastructure for Real-Time Inventory and Logistics in Retail Supply Chains. Educational Administration: Theory and Practice.

26. Johnson, A. E. W., et al. (2016). MIMIC-III database. Scientific Data.

27. Emerging Role of Agentic AI in Designing Autonomous Data Products for Retirement and Group Insurance Platforms. (2024). MSW Management Journal, 34(2), 1464-1474.

28. Kehl, K. L., et al. (2019). NLP for oncology outcomes. JAMA Oncology.

29. Segireddy, A. R. (2024). Machine Learning-Driven Anomaly Detection in CI/CD Pipelines for Financial Applications. Journal of Computational Analysis and Applications, 33(8).

30. McMahan, H. B., et al. (2017). Federated learning. AISTATS.

31. Varri, D. B. S. (2024). Adaptive and Autonomous Security Frameworks Using Generative AI for Cloud Ecosystems. Available at SSRN 5774785.

32. NIST. (2020). SP 800-53 Security Controls.

33. Singireddy, J. (2024). AI-Enhanced Tax Preparation and Filing: Automating Complex Regulatory Compliance. European Data Science Journal (EDSJ) p-ISSN 3050-9572 en e-ISSN 3050-9580, 2(1).

34. Price, W. N., & Cohen, I. G. (2019). Privacy and big data. Nature Medicine.

35. Keerthi Amistapuram. (2024). Federated Learning for Cross-Carrier Insurance Fraud Detection: Secure Multi-Institutional Collaboration. Journal of Computational Analysis and Applications (JoCAAA), 33(08), 6727–6738. Retrieved from https://www.eudoxuspress.com/index.php/pub/article/view/3934

36. Ribeiro, M. T., et al. (2016). Model explanations. KDD.

37. Varri, D. B. S. (2023). Advanced Threat Intelligence Modeling for Proactive Cyber Defense Systems. Available at SSRN 5774926.

38. Topol, E. (2019). High-performance medicine. Nature Medicine.

39. Paleti, S. (2024). Transforming Financial Risk Management with AI and Data Engineering in the Modern Banking Sector. American Journal of Analytics and Artificial Intelligence (ajaai) with ISSN 3067-283X, 2(1).

40. Hersh, W. (2015). Information retrieval in medicine. Springer.

41. Garapati, R. S. (2023). Optimizing Energy Consumption in Smart Build-ings Through Web-Integrated AI and Cloud-Driven Control Systems.

42. Bishop, C. (2006). Pattern Recognition and Machine Learning. Springer.

43. Inala, R. Revolutionizing Customer Master Data in Insurance Technology Platforms: An AI and MDM Architecture Perspective.

44. Jensen, P. B. (2012). Mining electronic health records. Nature Reviews Genetics.

45. Varri, D. B. S. (2022). A Framework for Cloud-Integrated Database Hardening in Hybrid AWS-Azure Environments: Security Posture Automation Through Wiz-Driven Insights. International Journal of Scientific Research and Modern Technology, 1(12), 216-226.

46. Chen, J. H. (2019). AI in clinical workflow. JAMA.

47. Silver, D. (2016). Mastering Go with deep RL. Nature.

48. LeCun, Y. (2015). Deep learning review. Nature.

49. Amistapuram, K. (2024). Generative AI in Insurance: Automating Claims Documentation and Customer Communication. Turkish Journal of Computer and Mathematics Education (TURCOMAT), 15(3), 461–475. https://doi.org/10.61841/turcomat.v15i3.15474

50. Raghupathi, W. (2014). Big data analytics in healthcare. Health Information Science.

51. Guntupalli, R. (2024). Enhancing Cloud Security with AI: A Deep Learning Approach to Identify and Prevent Cyberattacks in Multi-Tenant Environments. Available at SSRN 5329132.

52. Zhou, L. (2019). Data quality in health IT. JAMIA.

53. Smith, J. (2017). Hospital data lakes. IEEE Access.

54. Patel, V. (2018). Interoperability challenges. Health Affairs.

55. Koppolu, H. K. R., & Sheelam, G. K. (2024). Machine Learning-Driven Optimization in 6G Telecommunications: The Role of Intelligent Wireless and Semiconductor Innovation. Global Research Development (GRD) ISSN: 2455-5703, 9(12).

56. Dean, J. (2012). Large-scale ML systems. CACM.

57. Lahari Pandiri, "AI-Powered Fraud Detection Systems in Professional and Contractors Insurance Claims," International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering (IJIREEICE), DOI 10.17148/IJIREEICE.2024.121206.

58. Berg, B. (2013). Medical imaging informatics. Springer.

59. Rongali, S. K. (2023). Explainable Artificial Intelligence (XAI) Framework for Transparent Clinical Decision Support Systems. International Journal of Medical Toxicology and Legal Medicine, 26(3), 22-31.

60. Xiao, C. (2018). Opportunities in healthcare AI. IEEE Intelligent Systems.

61. Yu, K. (2018). Federated EHR analytics. AMIA.

62. Sun, J. (2017). Clinical time series modeling. KDD Health.

63. Inala, R. AI-Powered Investment Decision Support Systems: Building Smart Data Products with Embedded Governance Controls.

64. Gulshan, V. (2016). Diabetic retinopathy detection. JAMA.

65. A Scalable Web Platform for AI-Augmented Software Deployment in Automotive Edge Devices via Cloud Services. (2024). American Advanced Journal for Emerging Disciplinaries (AAJED) ISSN: 3067-4190, 2(1).

66. Char, D. (2018). Ethics of clinical AI. NEJM.

67. Mashetty, S., Challa, S. R., ADUSUPALLI, B., Singireddy, J., & Paleti, S. (2024). Intelligent Technologies for Modern Financial Ecosystems: Transforming Housing Finance, Risk Management, and Advisory Services Through Advanced Analytics and Secure Cloud Solutions. Risk Management, and Advisory Services Through Advanced Analytics and Secure Cloud Solutions (December 12, 2024).

68. Lee, C. (2019). Smart hospital architecture. IEEE Internet of Things Journal.

69. Rongali, S. K., & Kumar Kakarala, M. R. (2024). Existing challenges in ethical AI: Addressing algorithmic bias, transparency, accountability and regulatory compliance.

70. Breiman, L. (2001). Random forests. Machine Learning Journal.

71. Guntupalli, R. (2024). AI-Powered Infrastructure Management in Cloud Computing: Automating Security Compliance and Performance Monitoring. Available at SSRN 5329147.

72. Vapnik, V. (1998). Statistical Learning Theory. Wiley.

73. Nagubandi, A. R. (2023). Advanced Multi-Agent AI Systems for Autonomous Reconciliation Across Enterprise Multi-Counterparty Derivatives, Collateral, and Accounting Platforms. International Journal of Finance (IJFIN)-ABDC Journal Quality List, 36(6), 653-674.

74. Davenport, T. (2018). Analytics in healthcare. Harvard Business Review.

75 Keerthi Amistapuram. (2023). Privacy-Preserving Machine Learning Models for Sensitive Customer Data in Insurance Systems. Educational Administration: Theory and Practice, 29(4), 5950–5958. https://doi.org/10.53555/kuey.v29i4.10965

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

76. Snow, C. (2017). Data governance frameworks. Information Systems.

77 Chava, K. (2024). The Role of Cloud Computing in Accelerating AI-Driven Innovations in Healthcare Systems. European Advanced Journal for Emerging Technologies (EAJET)-p-ISSN 3050-9734 en e-ISSN 3050-9742, 2(1).

78. Stead, W. (2017). Clinical data standards. JAMIA.

79. Rongali, S. K. (2024). Federated and Generative AI Models for Secure, Cross-Institutional Healthcare Data Interoperability. Journal of Neonatal Surgery, 13(1), 1683-1694.

80. Mandel, J. (2016). SMART on FHIR. JAMIA.

81. Benson, T. (2012). Principles of health interoperability. Springer.

82. AI and ML-Driven Optimization of Telecom Routers for Secure and Scalable Broadband Networks. (2024). MSW Management Journal, 34(2), 1145-1160.

83. Marx, V. (2013). Data mining in medicine. Nature Methods.