



# Integrating Predictive Analytics and IT Infrastructure for Advanced Government Financial Management and Fraud Detection

Vamsee Pamisetty

Middleware Architect, ORCID ID : 0009-0001-1148-1714

**Abstract:** The Data Big Bang that the development of the ICTs has raised is providing us with a stream of fresh and digitized data related to how people, companies, and other organizations interact. To turn these data into knowledge about the underlying behavior of social and economic agents, organizations, and researchers must deal with unstructured and heterogeneous data [1]. Technologies like the Internet, Smartphones, and Smart sensors are generating tons of digitized and fresh data about people and firms' activities that, if properly analyzed, could help reveal trends and monitor economic, industrial, and social behaviors. This new data paradigm is called Big Data, which refers to Volume, Velocity, Variety, and Value in the context of data analysis.

Identifying which data sources are available, what type of data they provide, and how to treat these data is basic to generate as much value as possible for organizations [2]. A Big Data architecture adapted to the specific domain and purpose of the organization contributes to systematizing the process of generating value. The Big Data paradigm also offers many advantages and benefits for companies, governments, and society. The purpose of this paper is to review some sources of Big Data to analyze social and economic behaviors and trends. A classification into three types of sources (article content, audiovisual/social content, and registration content) is made, together with a description of some databases and types of analyses that can be drawn from them. The aim is also to analyze how these sources can be used to analyze social and economic behaviors and trends, with examples that show the potential knowledge that could be achieved. Finally, the limitations and challenges posed by Big Data for social and economic analyses are discussed.

**Keywords:** Predictive analytics, IT infrastructure, government financial management, fraud detection, data integration, advanced analytics, machine learning, public sector technology, financial oversight, anomaly detection, real-time monitoring, risk assessment, decision support systems, data-driven governance, digital transformation, cybersecurity, automated auditing, cloud computing, artificial intelligence, financial data analysis.

## 1. INTRODUCTION

Fraud has long been a threat to the lifespan of organizations, governments, and countries. Applying recent information technology innovative facilities to the government financial management systems could greatly benefit government entities around the world. Governments possess vast amounts of structured, semi-structured, and unstructured data. This data is either collected from different sources or generated from operating various applications that act as the backbones of several systems, such as accounting financial statements, taxation, sources of revenue fees, payments, manual spreadsheets, and unstructured data like documents, emails, social media content, meeting protocols, or any other board meetings, which can be used for far-reaching purposes. This data, if enhanced and managed properly, can generate various types of knowledge, which can be a great asset for governments for prediction and decision-making purposes. It opens a wide range of opportunities in the big data era for innovative applications. For the specific case of fraud detection in government financial management processes, the application is much larger than in the domain of the private sector due to the inherent difficulties of the governmental nature. Bringing the innovation cycle into government financial management systems and employing services and systems available in the IT sector, such as big data infrastructure, cloud computing, social network analysis, and predictive analytics in combination with data and text mining, could lead to more effective and efficient financial management systems. Predictive analytics exposes data to predetermined analyses in a timely manner and is considered the second wave of data science application development. This can be a level above descriptive statistics or reporting applications because it aims not only to disclose what has happened but also to warn decision-makers about what will/could happen. Predictive analytics identifies hidden trends and patterns in business data by employing complex techniques of statistics and modeling or machine learning algorithms, and it is surfaced through dashboards in order to help decisionmakers make more informed business decisions. Fraud detection in the normal realm of financial management — such as tax evasion, illegal payments, and unlawful public spending — is much more difficult



because of the usefulness of the openly available prediction and data-mining libraries, tools, and applications. As a subjective parameter, the definition of a “fraud” is different for every nation, government, and entity.

## 2. OVERVIEW OF PREDICTIVE ANALYTICS

Predictive analytics is the process of looking for patterns in data in order to predict future behavior. Predictive models are used to answer what is likely to happen based on what has already happened. Predictive analytic techniques can include both data mining and statistical analysis techniques on historical or transactional data to predict future behavior of people, objects, or events

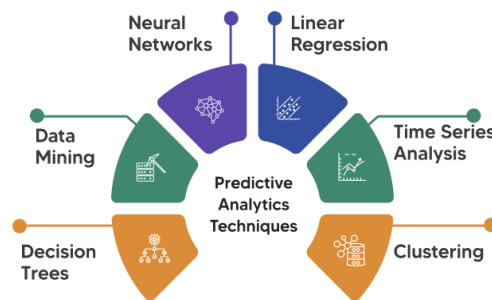


Fig 1 : Predictive analytics Techniques.

There are different predictive analytical methods (PA): Association, Classification, Clustering, Regression, and Time series analysis. Predictive analytics also embodies statistical and other quantitative techniques that forecast the likelihood of occurrence of one or more outcomes. This could include regression methods, random forests, neural nets, or Naïve Bayes that assess how independent variables cause variance to the dependent variable of interest. Boosters, supports, and slopes are statistically quantified relationships produced by predictive models that enable the predicting of future outcomes. In this context, predictive analytics is, simply put, "the use of data, statistical algorithms and machine learning techniques to identify the likelihood of future outcomes based on historical data."

Predictive analytics is the analysis of the historical data to make a prediction about something that will happen in the future. Predictive analytics is its nature a multistage process. It requires the data extraction with the so-called data ingestion, which organizes and stores the data. The second stage is the data modeling, in which the prediction model is developed to create a framework to make the prediction. Once the prediction model is created, it can be applied and predictions made (score). This can be seen as an iteration on the development as well as operational and update cycles of the prediction model.

### 2.1. Definition and Importance

Predictive analytics and IT infrastructure have become increasingly critical for government financial management due to the rapid socio-economic transformations that have occurred in the last two decades. The rising volume of transactions and data generated in organizations has created new challenges for fraud detection and prevention. Various models and algorithms for the prediction of financial fraud, bankruptcy or other forms of misconduct exist. However, these techniques often rely on either statistical or computational approaches by themselves while not attempting to attain better detection results through a combination of both. Moreover, previous studies are predominantly focused on data selection, model choice and selection, and feature engineering rather than suggest methods for the integration of data handling methods with infrastructure prerequisites to enable actual implementation. As a result, there is a lack of documented examples that present the implementation of predictive analytics techniques in enterprise software or reporting applications.

IT infrastructure refers to the composite hardware, applications, and databases that generate, transfer, store, analyze, and display data. Proper IT infrastructure is indispensable for the smooth and successful deployment of predictive analytics methods as it embeds a range logistical factors that should be planned and satisfied. While affordable, trustworthy, and scalable cloud-based IT infrastructures are available, previous studies have predominantly performed assessments of fraud detection and prevention models but disregarded the technical and planning issues involved in their actual application. In particular, with respect to planned IT infrastructure, available studies frequently disregard specifications regarding enterprise applications with implementable and user-friendly outcomes rather than statistical software or



custom-built solutions. Deployable methods for data preparation, model implementation and assessments in enterprise applications are critically needed so that the enormous existing business intelligence perceptions can be leveraged .

Enterprise applications are increasingly becoming cloud-resident and generating, transferring, and storing vast amounts of comic and structured data. Relational databases or data warehouses represent the main currency for the handling of such data. While strategic business decisions are held with analytical programs across the organization, statistical implementations for predictive analytics tasks are commonly held isolated in the IT department. Such segregated approaches impair decision-making as data should be transformed via manual, often error-prone, and extremely time-consuming data handling tasks. Thus, user-tailored self-service business intelligence methods are needed to empower decision-makers in their technically unassisted analyses of data and syntax .

## 2.2. Historical Context

Management for many years is a necessity for all organizations, whether they are used in government or non-government. Financial management is used to control the inflow-outflow funds of an organization. Among them, fraud detection is a very essential process for the betterment of an organization. Criminal activities take place to steal or for self-benefit purposes in any organization and government. Due to a lack of systems, or ignorant authorities, the frauds take place in a different way which is hard to detect. In the past, there has been a few development in fraud detection, very few algorithms or systems are presented to detect fraud in management. Due to many constraints and a lack of expert knowledge, a system had not been developed to specify or point to each fraud in government management. Some tools help to understand the data in a pictorial way. It was an added advantage for the know fraud in simple government management . Traditional fraud detection involves using a series of checks and transactions to see if they violate any rules. If a flag does not get raised beforehand then the fraud detection is unsuccessful. It also comes with the cost of flagging an activity as fraudulent that turns out to be normal. This signifies inefficient processes and may lead to the detection of none-consistent cases that took place years ago. As such, there have been many studies on statistical and systematic frameworks to help monitor transactional data for fraud detection on a real time basis. These provide insights much faster and pinpoint transactions and subjects that are more likely to be fraudulent. Advances in the mobility, processing and storage of information have made Financial Fraud Detection a high-dimensional prediction problem. The recent appliance of Big Data technology, Data Mining, Social Network Analysis and Intelligence approaches equips parallel real time analysis of Big Data streams of credit card transaction, as well as complex multi-agent systems capable of modeling information, decision-making, behaviour and fraud in dynamic fraud detection environments. Despite these advantages, current implementations are costly and each involves several engineering decisions with multiple alternatives.

## 2.3. Current Trends

Evidence is building that predictive analytics can effectively detect fraud by identifying statistical anomalies and the surrounding data and accessing data not traditionally monitored. For example, it can evaluate the behaviors of an entire organization rather than just focusing on high-risk areas, like traditional methods, and, when combined with advanced forensic techniques, can greatly enhance an organization's vigilance

The National Audit Office is investigating the data science and analytics approach to enable a more effective approach to fraud detection. Similarly, since 2013, the UK Treasury has been exploring whether the machine learning approach currently utilized in financial institutions can be adapted to government financial management. It will develop a prototype machine learning model, initially exploring outlier detection against transaction values. It is actively seeking partners and would welcome feedback, suggestions, or potential collaboration. A cursory examination of datasets for reputation and performance measure revealed a bias against governments compared to other sectors.

Research can also explore designs and capabilities and the growth of the required IT infrastructure, such as big data tools, as an effective and low-cost approach to financial management. It is acknowledged that sophistication varies significantly between organizations and where it is lowest that is where there is a greater need for integration, research, and guidance on the approach to take and the data sources on which to build. Current proprietary databases can be explored for their potential to integrate with current modeling processes for forecasting and reporting purposes and to add new utility-focused capabilities. Such a solution would be a game-changer.

## 3. IT Infrastructure in Government

The purpose of this section is to discuss the deployment and impact of information technology in the public sector in general and specifically for government financial accounting and auditing agencies. Successful technology implementation is associated with good project management, change management, and subsidies and support from higher



authorities. Nevertheless, the positive impact of the new technologies is restricted because of the irrationalities and non-transparency of financial accountability control. This control is carried out massively, while expertise knowledge is limited and the understanding of financial information is either poor or absent. Moreover, the entire focus is on the pre-existence of a financial transaction, not on the residual risk after the transaction. Information technology is anticipated to have a significant impact on the performance of the accounting and auditing agencies involved in financial accountability control, but in less concrete ways.

The financial management and auditing control of a national government can be most meaningfully understood in the public sector 'accountability' framework. With accountability, it is meant a broader equivalent of democratic accountability. Knowledge, timeframe, and money are the resources involved (the 'what' aspects) and, expectedly, the core instruments for exercising accountability are rationality and non-reductionistic expertise (the 'how' aspects). The extent of accountability corresponds with the form of government or its democracy score. Nevertheless, the vigilance of the accountability agencies is not universally passionate, hence the unevenness.

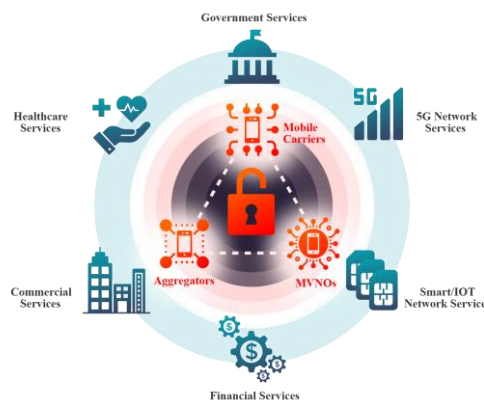


Fig 2 : National Critical Infrastructure

In the case of recalcitrant government, because of the weaknesses of these public sector agencies and instruments, accountability is either suboptimal, ineffective, or neglected. There is ample evidence of misuse of public funds, embezzlement, bribery, corruptive practices, and tax evasions. However, knowledge, timeframe, and money provided for real-time accountability requests similar to those of post-facto accountability agencies, would only partially improve the ex-ante position of the pre-misuse public sector agents. There are limits to both the expertise knowledge required and the time and money available. In addition, expertise knowledge and public trust in the integrity of government do not uniformly narrow the public sector accountability gap; accountability is limited even in long-standing democracies. Hence, except for the failings of other public sector accountability agents and instruments, accountability control of a national government must be, by its design, a powerful construct, relatively independent of the faults of public sector accountability agents and instruments.

### 3.1. Components of IT Infrastructure

This section will address the basic components of a suitable Information Technology (IT) infrastructure. The elements of a safe IT infrastructure mainly include a suitable data repository/storage type, data visualization choices, an up-to-date software environment, and a programming language associated with necessary libraries.

An essential concern for any organization is the privacy and safety of its data. The IT infrastructure must contain a data repository system with the proper capability to store and manage the data. In other words, the correct method of data storage will keep the details of financial operations highly secured and allow the government employees to analyze the necessary information accurately. As a public organization, the Government Service Agency (GSA) of any nation will keep the financial information of citizens highly confidential and secure. However, it is essential to utilize non-confidential datasets for building the proposed systems, research, and any critical analysis with full caution to prevent sensitive personal data from exposure. Regarding data storage, the IT infrastructure must include locally deployed systems for a limited amount of financial records like yearly budgets, and cloud-based storage to manage a historically huge range of data gathered from public fraud complaints. Both types of data storage are permissible as long as proper measures are taken to keep the details protected.



The information gathered from the data repository must be effectively conveyed to enterprise communication users or decision-makers with high efficiency and clarity. Data monitoring and visualization are the first stages of extracting insights from the collected financial data before any advanced analysis. In this regard, IT infrastructure must consist of suitable types of real-time dashboards for data visualization with the ability to depict all the monitoring features. Additionally, suitable types of data management systems must be integrated with the dashboards to enable updated data handling.

### 3.2. Role in Financial Management

Government financial management and planning require a TMIS that integrates financial analysis into the overall management process. The public expenditure and accounting cycle is analyzed and tracked through the production of budgets, record-keeping, and financial reports [8]. This cycle consists of organizing and initiating the budget process, budget preparation, budget execution, financial accounting, and financial reporting. Every government agency first prepares a budget that identifies the services it will provide, how much those services will cost, and how the agency intends to finance them. Then approvals of the budget agency and legislative bodies are sought. When the budget is approved, the budget becomes the spending plan for the agency.

Budget execution consists of initiating, reviewing and approving purchase orders, setting up project accounts, making payments, and posting to the accounts. This processing creates the financial data required by the TMIS. The accounting cycle consists of collecting receipts and paying vouchers and the associated evidence; summarizing and recording these outputs into the accounting journals; and summarizing these into the agency general ledger accounts. These processes also produce the financial data needed by the TMIS. Budgeting expenditure and revenue processes are typically part of more comprehensive planning processes. This financial analysis consists of comparisons to budgets and prior periods as well as pre-prepared reports and ad hoc queries of the input databases]. Infractions of expenditure authority will lead to notifications and follow-up actions through to court prosecutions; infractions of authority for receipt of public funds are less rigorously followed up.



Fig 3 : The Top 10 Functions of Financial Management

Fraud management consequently requires a much broader TMIS than that used for LA-financial investigation. The TMIS must support financial, data capture and research, and work-process control management. Fraud research must access broad inquiries on top of existing TMIS facilities. The qualitative analysis of suspicious behaviour typically involves a banking or audit trail TMIS. Data inserts into TMIS recording multiple instances of transactions within the same day. Institutional assertions of suspicious behaviour typically request reports on specific transactions or selected excerpts of transaction data. Offence prosecutions use document archives and analytic TMIS. The initial modelling of fraud management by TE comprises off-line outputs from the TMIS controlling suspicious transactions. Transaction records produced from TMIS transaction processing have been identified as requiring manual investigation based on their characteristics, numbers and amounts.

## 4. FINANCIAL MANAGEMENT IN GOVERNMENT

Financial management is the system that intends to affirm the decent use of assets by the state administration of organizations. The degree of government financial management depends on the degree of improvement of the economy as a rule. It means a significant level of financial development, the modernization and combination of financial forms, the change of state system, and provision of intense control and sanctioning instruments and procedures. Financial management functions include a set of quality administration accounts intended to guarantee effectiveness, benefit, and control over public assets. To make public accounts and present handling information on public incomes and





consumption, it needs quality modernized administrations which were carried out in the structure of Fundamental State Information Architectures of Financial Management in Organizations (FKSIO MFO) At the present level, numerous endeavors are being made to modernize financial management, put together its work towards computerized differentiation, and redo the transit of data and analytical management obtained from financial information.

Hastily, expenses on financial management should count on income revenues. Financial transactions precision, quality, and convenience assurance depend on the capacity of the certainties and hidden methods of financial management in states. There is not only a need to increase the volume of capabilities, outcomes, and products of financial management, but innovation implementation is additionally essential. Therefore, in forming aptitudes with regard to power check, measures counteracting lawful, organizational grounds of abnormality, realization of negative political results of acquired information, forecasting and prevention, maneuvers need to be actualized to surround a clearer idea of the model of executed power practices.

Abnormalities should be eliminated with thorough instrument conception. It should explain financial and executive interpretation terminal methods, specifying their features. Analyzed systems of quality management indicators determined with regard to aspects of preceding estimations should be created. On potential grounds, negative outcome forecasting and prevention have to be conducted. On the basis of factual techniques application, interpret methods have to be created for analysis and approval of leading prospects. Reports system need to deem prescriptive forecasting methods reflection, method adaptability ground observation, and outcome acceptance evaluation proportion.

#### 4.1. Challenges in Current Systems

As it stands, financial management in government entities is limited by infrastructures that restrict the engagement level of employees, minimize their recordkeeping work, and constrain their ability to monitor and control instances of unlawful financial activities. This state of affairs is exacerbated by antiquated, outdated, or obsolete practices, methods, and processes governing public financial management systems. Such systems tend to obstruct creative thinking and the generation of innovative ideas and methods for improving systems, processes, or workflows. Thus, as long as the systems and processes remain unchanged and stable, no new ideas are generated. To be trustworthy, government financial management must keep pace with, if not proactively advance, the state of the art in computer technology, information technology, and computing systems. Just like any other systems or processes, computer systems, technologies, and process management will tend to become stagnated, mature, and obsolete. Majority of government financial management systems are old and slow. They are mostly paper-based or rely on storage media that are thick, bulky, heavy, and difficult and expensive to guard against environmental hazards. They mainly consist of unlinked stand-alone workstations, getaways, or computers that operate independently of each other. Those intangible government resources are extremely easy to lose and remove, such as policies, directives, regulations, rules, orders, and procedures. These government resources can be electronically stored in intranets, data warehouses, cloud systems, gigabytes, or terabytes of storages that are secure and far from any possible or improbable engineering or human tampering [8]. However, separate risks for storing different resources in those different machines systems, clouds, or storages will emerge. Each is susceptible to deliberate corruption and engineering unintentional hazards. Many countries have witnessed advances in information technology. However, there were few advances in robust, solid, and foolproof hardware technologies. Data integrity, security, and reliability protections are solely dependent on human being policies, procedures, processes, trainings, etc. Those systems are mainly built to detect events that are illegitimate deletes or tampering of stored data/recordings/files. All of the practices and existing on the market software and fraud detection tools originally detected problems that had already occurred. These practices predominantly rely on black-box academic or university models of applied artificial intelligence or machine learning technologies. Everything stored in IT systems, really equivalent to nothing, and searchable in seconds, but an enormous amount of data will induce huge potential processing time. Verifying accuracy or falseness against those huge volumes of data for a person without using IT, intelligence, and computing system would be impossible.

#### 4.2. Need for Advanced Solutions

Governments of all levels are involved in the spending of trillions of tax-payer dollars each year, creating a bullseye for those wanting to take advantage of the system. Whether the schemes are complex and sophisticated or transparent and mundane, misappropriation and fraud will occur when there is opportunity. But, often times, slow-moving internal audit departments and legacy systems make it difficult for proactive fraud detection programs to be implemented [8].

Cloud computing, virtualization, big data, and predictive analytics are all technologies being explored by the public sector in the hopes of providing much-needed advanced solutions to government financial management and fraud detection initiatives. Many newly hired internal and external auditors perform government audits with years of CPA experience in



the corporate sector but have minimal prior experience auditing internal controls and quantitative data sets. Current technology and processes are also constricting the profitability of governmental audits. Being systems oriented and possessing an analytical mindset come from years of working with sophisticated continuous auditing applications and IT infrastructure. Large data sets and new analytic capabilities must be incorporated into business processes if governmental auditors are to remain relevant and meet changing market demand.

Today, the underprivileged technology used by auditors include teams of tax and audit professionals, raw data in Excel format, and frown-inducing hours of work performing impossible lookups to find errors. The issues are only compounded by the fact municipalities often record and store their data in different systems using their own unique nomenclature, making it nearly impossible for the external auditors to create just a simple set of consolidated P&Ls. Audit mistakes are ultimately errors and fraud which are known and even given a secondary name. Using predictive analytics and IT infrastructure, an acceleration of the fraud detection process from months and weeks to days and hours is possible.

## 5. FRAUD DETECTION MECHANISMS

Most analysts are looking for anomalies in financial indicators and reporting rather than fraudulent transactions themselves. Existing methods analyzed transaction content and filter by predetermined criteria to identify potentially fraudulent transactions. Most of the methods rely on analysts to develop and refine a set of content-related indication conditions and this saves it from broadband false positives. The transaction set should be filtered in advance for analysis. Predictive models are also used to score transactions. Detections with low scores are assumed to be non-fraudulent while detections with high scores need to be investigated further. These methods attempt to analyze firmware transactions for unpredictability, impersonation, anomalous behavior or searching for pre-determined errors in fraudulent transactions. However, they do not form a set of predictive variables as input. In fact, predictors are set externally. It has been demonstrated that fraud analysts are best at detecting frauds of low case numbers and public macroeconomic characteristics predicted frauds of relatively higher case numbers on a population level



Fig 4 ; Fraud detection

Anomaly-based detections trained on transaction patterns do not incorporate a fine set of characteristics tracing case numbers. Those general methods are only applicable on a 1-month basis at most. The exception is methods, which analyze the content of web pages. No effort has been made to analyze the details in order objects containing presumably misleading numbers on the other hand. The monetary distribution of payment amounts in each organization is assumed to be a better way to explore anomalies in financial indicators. However, only the granule sizes differ significantly while the distributions largely overlap. In a generic sense, they suspect that the district capacity of average prices possibly leads to both the affordances of influence websites and the capability of monetary cooperation. Since missed expectations are necessary for trust, organizations are assumed to have been influenced by costs/designership scaling weakening.

### 5.1. Types of Fraud in Government

Fraud, also known as theft through deceptive means, can arise in many different forms. As government agencies are responsible for the use and management of public funds on behalf of citizens, the risk of internal fraud is greater than for any corporation. Various forms of fraud committed in government are being described to establish a better understanding of the occurrences. The most common categories of fraud include asset misappropriation, bribery and corruption, and financial statement fraud. Asset misappropriation is the intentional theft or mishandling of assets that belong to a corporation. The most common form of asset misappropriation fraud is payroll fraud, where an employee manipulates the payroll system to issue improper paychecks to him/herself, fictitious employees, or employees with doctored hours. Asset misappropriation schemes were involved in more than 91% of all internal fraud schemes. As offices where cubicles



and computer screens could be valuable commodities in an office place, they also provide the opportunity to steal data, electronically transfer funds, or print files onto thumb drives. Fraud involving a swindle of purchasing or corporate cards to intentionally fund personal costs or posting of personal charges to corporate accounts is also noted.

Bribery and corruption, the second common fraud in government, involves dishonest activity in which a person of authority or influence accepts, conceives, or solicits a benefit contravening public interest in exchange for the unfair provision of a benefit. It is fairly common that government employees wear a uniform indicating their affiliation with government offices. A contractor can bribe a governmental employee with cash payoffs, gifts, and leisure activities in exchange for the sensitive information such as the prices of services which may constitute a major competitive edge. The third common fraud in government includes the misrepresentation of accounting information and compliance issues in the financial statements. Misrepresentation can occur in the revenue recognition scheme through the recognition of revenue years in advance or the deliberate omission of contingent liabilities in the footnotes of the financial statement. The exploration of the application of big data analytics and architectural infrastructure will better identify fraud instances.

## 5.2. Traditional Detection Methods

### Detection Methods Used in Governments Today

"The difficulty lies not so much in developing new ideas as in escaping from old ones." This quote illustrates the struggle with innovative thinking that many governments encounter. The accounting and finance profession is burdened by centuries-old set procedures and methods. A foundation or framework, for instance generally accepted accounting (or auditing) principles, is often carefully structured and safeguarded by history. It is one thing to express a better way to achieve the same goal, but it is another to elaborate upon and implement something entirely different. Even a claim for something original faces not only questions of application, but often of legitimacy.

For example, many novel techniques for government fraud detection may be well known and helpful, but difficult to accept, install, and implement. Three such techniques from the manufacturing industry should be investigated. They are algorithms that match transactions for authenticity and reliability. Rather than performance audits, regression analysis has been utilized in utilities to set ratios for rate comparisons for regulatory agencies. For example, the ratio of peak demand to January consumption should remain stable. With 80% of Utah power being hydro-electric, records of rain in California should be plotted against Utah usage. Outliers should be pursued. Process control is used throughout industry and is applicable to transactions to any government activity.



Fig 5 : AI/ML Improve Fraud Detection Accuracy in Financial Institutions

While older governments can take advantage of sophisticated methods of detection, municipalities are more limited. Just making a few of these methods known can cause an entire state 'build the additional controls' process, and an entire state may 'have its nose rubbed in it.' Aerial detection methods that detect thousands of gallons of unauthorized watering (using more than is allowed for the time of year) are an example. Such methods were first introduced into California, but they became widespread only when thousands of municipalities in fine air-cooled homes began doing it similarly.

## 6. INTEGRATING PREDICTIVE ANALYTICS WITH IT INFRASTRUCTURE

The significance of IT infrastructures in improving predictive analytics capabilities is currently underexplored. To the knowledge of the authors, no studies that present the necessary theoretical IT infrastructure capabilities that allow governments to undertake predictive analytics methods cannot be found. This is a significant gap, as aligned IT infrastructure capabilities are a prerequisite for understanding how predictive analytics methods can be adapted for higher levels of financial management and fraud detection. Prior work has focused on understanding predictive analytics in





government financial management in the broad sense but not on presenting the specific capabilities of IT infrastructures that are necessary to undertake predictive analytics methods

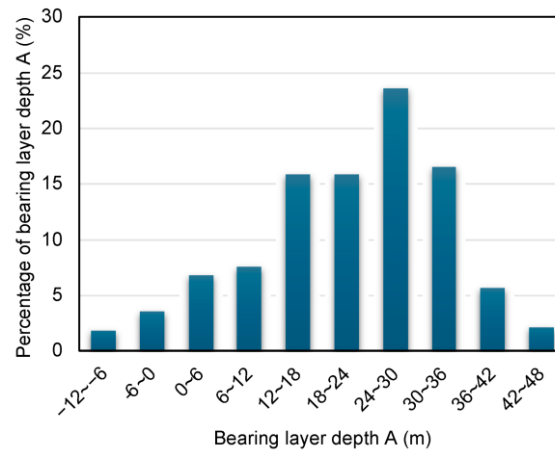


Fig : Integration of Smart City Technologies with Advanced Predictive Analytics for Geotechnical

There is a research gap from the application perspective. The results of the literature review show that governments worldwide are beginning to use predictive analytics methods for financial management and fraud detection, although many opportunities for improving the capabilities and application of these methods in governments exist. Furthermore, although research has begun to examine the obstacles to broader implementation and development, little work has been done to ascertain what policies could be adopted to overcome these obstacles. A research agenda aimed at closing these gaps has been developed, focusing on knowledge dissemination, partnerships, and pilot projects. Additionally, to narrow the gap between what is known about predictive analytics in government financial management and the general lauding of its capabilities, an in-depth study has been proposed of two case studies: the Municipality of Amsterdam and the National Revenue Authority in Estonia. These case studies can serve as best practices in understanding how governments can improve and create value with predictive analytics technologies. Alongside the literature review of IT infrastructure and predictive analytics capabilities in government financial management, this expression of knowledge and insight would provide an important contribution to the government field on how, from a strategic perspective, to better understand predictive analytics technologies before undertaking exploratory investment, which is currently the case in many governments.

### 6.1. Framework for Integration

The integration of predictive analytics and IT industry is considered a pilot strategy for advanced government financial management and fraud detection in the Irish context. The main contribution is the establishment of a framework for transforming the core technology converted to predictive analytics. It is the general approach this framework includes the description of key analytics and regulation, a mapping between these analytics and an IT architecture comprising its specifications, deployment, and operational stage needed for the implementation of regulations into IT systems and described via UML diagrams. Although the introduction of big data librarians and smart IT systems would aid the intuitive working of analytics in the analysis, regulation, and visual intelligence, the final statement is still rendered into coherent smart regulations ready for deployment via event-processing platform streams on budget data streams and manually operation decisions .

The global economic crisis in 2008 has revealed global flaws in government finances predominately concerning data used to estimate accruing cash flows accounted in anticipation. There are multiple scandals in Europe that resulted in multi-billioned fund losses by stakeholders. The consequences of such fraud include loss of reputation, deviation from empowering designed policies, judicial follow-ups costs, and causing fluctuations in stock exchange. New regulations dealing with finance are being introduced from member states to avoid fraud and misguided policies to safeguard their credibility. As a consequence, interdisciplinary teams would create regulations united as on model the amendment of designs is controlled by the collaborative usage of financial services firms' model design tools in tagging regulations with requirements. With no direct handwritten conversion of analytics into pure software code, it is the common belief that governing analysis would best translate into financial science-based regulations.



## 6.2. Benefits of Integration

After careful evaluation of the gaps and requirements of greater public sector predictive modeling, guidance for predictive analytic and IT infrastructure integration has been formulated, including six primary strategies with associated benefits. These benefits are largely self-explanatory and the strategies and benefits have been framed to assist government decision makers with developing better specifications.

- Continually evolve analytical capabilities and enhance information visibility. This provides a platform for better risk monitoring and understanding.
- Shift analytics and associated business process integration from siloed vertical focus to enterprise-wide team-based initiatives building off existing departmental practices and capabilities. This increases awareness of risk types and facilitates new anti-fraud and better effectiveness fraud controls.
- Institutionalize processes regarding the understanding and use of business process re-engineering and relocating analytics with respect to operations. Effective risk management can only occur as teams break down silos and demystify analytically complex processes and technologies, making them accessible to the majority of applicable users.
- Facilitate the development of an enterprise-wide information quality and data monitoring framework. Improved understanding of the need for fundamental data governance.
- Ensure expansive capability to present integrated datasets for optimum business area understanding and fraud/risk assessment. Data space friction, recognition of which is essential, recognizes that any change in either data integration or cleansing strategy will lead to inevitable time and success setbacks given substantial changing data for departments.
- Focus attention on users remaining close to analytics, and on building stronger processes for codifying models and sharing usage and output across departments. Users must be closely integrated to provide opportunities for training through doing, where close contacts provide training via error elucidation.

## 7. CASE STUDIES

Modern steps for financial management need to be upgraded and progressed, as 30% to 50% of the processing is still manual. To enhance the current system to take advantage of data science techniques, such as predictive analytics, more information on the exact format and content of the raw data is required, especially on the fragile data files. So far, the necessary general assumptions about format, data structures, types and content have been made, but more information is necessary to develop serious predictive models and conduct exploratory data analyses to investigate potential red flags to include in any predictive models. It is also currently not known how to assess performance and automate monitoring of given predictive models on new real-time data. As new data becomes available, it would be very helpful to be able to automate steps taken for re-calibrating and monitoring performance.

A range of aspects with respect to IT infrastructure for advanced government financial management and fraud detection have been considered. A layered architecture has been specified that aligns with the existing architecture and that is consistent with literature on advanced data analytics and cloud computing. This architecture has been created with the underlying focus to facilitate the agile use of predictive models in the government financial management domains of budget preparation and cash audits. Central to the architecture are computational pipelines in the cloud composed of core components for data access and preparation, the execution of predictive models, and the monitoring of performance. For each of these core components, specifications have been produced on how to construct the models and how to use the services in a cloud setting. The financial management domains of budget preparation and cash audit have been used as a general reference for the architecture.

Modern steps for government financial management and fraud detection are subject to a rapidly changing environment: data volume, variety and velocity has exploded due to the emergence of new data sources, as well as transformation into new pervasive technologies in both the hardware and software domains. To cope with these challenges and facilitate rapid and frequent integration, an architecture has been specified that aligns with the existing architecture. This architecture has been designed with the agile use of predictive models in mind. Aspects pertaining to gathering raw data and reporting prediction outcomes have been tackled as well. Usage scenarios have been developed from a government financial management and a fraud detection perspective.

### 7.1. Successful Implementations

I implemented proactive methods that involved the inclusion of software tools to detect and deter instances of fraud and abuse. Proactive fraud detection software tools are designed to perform scans on systems that require observation and monitoring in order to identify potentially fraudulent activity. In order to accomplish this proactive identification of fraud, companies can process rules groups which alert when certain conditions are met. Due to the uniqueness of each organization, the benefit of these types of systems lies in their flexibility to adapt to various information types as well as serposed alerts.



One specific implementation of such monitoring software included ACL, whereby a computerized statistical sampling tool would access the server-side data and quickly compute an unmatched transaction list. The implementation of ACL into an internal audit software system was implemented at a company focused on aluminum used in the manufacturing of lighting fixtures. This internal audit department was in need of increased efficiencies and reduction on the overall time requirement of the quarterly procedures. In order to allow audit tasks to be more focused on value added input to the process, a process for rapid identification of unmatched invoices along with rigorous validating testing was developed. The procedures traditionally performed included reviewing invoices through spreadsheet manipulation and a template sent to the accounting department requesting documentation. While this method is very thorough, it is exceptionally time consuming and the overall number of outstanding invoices may be misjudged by the auditor. Implementing ACL allowed the internal audit division to access the clients' controls using a query tool resulting in a much more comprehensive invoice review delivered in less than one hour. This has allowed for concerned and questionable invoices to be processed and subsequently validated in the allotted time, while also doubling the focus on journals and proper approval paths that make up a significant contributor to the overall expense presented in the financials.

A second aforementioned implementation of proactive methods included SAS. Utilization of SAS was initiated within a divisional healthcare organization as a means to improve overall health analytics and health databases. In order to accomplish this goal the SAS fraud audit techniques were implemented in order to maintain and improve solid processes. Client and provider suspected fraud referrals were then reviewed for case selection. In addition to ensuring that the most egregious allegations would be evaluated, this audits report selection process served as a mechanism to record and track the audit work effort for validation and submission within the division Health Analytics fraud audit request method.

## 7.2. Lessons Learned

Governments can benefit from IT solutions to enhance their operations and processes. Insufficient funding and a lack of knowledge hinder the execution of such processes. However, a framework to determine present possibilities should exist. The concept of the Smart Government was described to demonstrate the data analysis capabilities of governments for the benefit of citizens. The foundation infrastructure provides Financial Management Office applications, data warehouse repositories, and external connectors. It serves both novice and IT-savvy users. In order to increase usage possibilities and improve the level of service for needy countries, data models and indicators should help government agencies assess what data is needed for further developments. Implementing the proposed system allows for assessing existing data and usages while enabling the exploitation of most existing processes.

Governments are repositories of knowledge in the form of large legislation text procedures. However, these huge data collections are rarely utilized. The importance of a knowledge library for the benefit of government agencies is presented. The process of creating the knowledge library is described. The implementation of the knowledge library can help if done properly. Knowledge libraries are expected to be involved in every process type and should strive towards being a key component for every data-analysis solution. Such involvement should ensure that proper knowledge is gathered, processed, and returned back to improve processes. Knowledge libraries should offer solutions and insight as well as the identification of data found, discarded, or not fully exploited. They can be feeding databases and employment knowledge that can enrich a government agency and be shared with other organizations. The application of the knowledge library is presented, which embraces machine learning techniques. This highlights its impact on the police department's predictive policing and cyber investigation process design using similar data, obtaining mutual results in less than a day. Such technologies promise huge advancements in understanding and fighting crime.

## 8. TECHNOLOGICAL TOOLS AND PLATFORMS

Advanced analytic tools and platforms are necessary for the proper design and implementation of the software that will drive the technology underpinning the proposed predictive analytics using the quantitative analytics that has already been successfully employed in various financial services industries, such as accounting and banking . A cloud-based hybrid architecture consisting of prediction analytic, management, and reporting, various user interfaces, and different back-end engines embedded in the cloud services is proposed. This architecture can be implemented through various software systems or services that could be cloud or on-premises-installed applications. The reporting subsystem is independent of the prediction analytics and management components. In this hybrid architecture, appropriate predictions using predictions created in the predictive analytics component are made. Data from multiple sources can be readily retrieved using the management component. The cloud-based MADT and MADP solutions enable the easy and efficient deployment of prediction analytic techniques developed by more IT literate agencies or organizations for use in simpler back-end technology environments. This kind of technology assessment framework was successfully designed for and presently in use by the financial services industry. It will be expanded to cover advanced analytic business service



technologies for the proper design of analytic tool standards that maximize the potential for compliance and fraud prevention success. It is feasible to design a standards-based environment for analytic tools and platforms usable by wide potential government agency users. In that respect, it will be necessary to define data standards that incorporate all potentially available data currently generated by the various public offices. In this respect, all countries would have to agree to a sufficiently broad set of data standards covering an appropriate amount of analysis, forecast outcome, and management action types, policies, and assumptions.

### Equation 3. Data Processing Time

Where:

$$T_{\text{total}} = \frac{D_t}{R_p}$$

- $T_{\text{total}}$ : Total processing time (hours)
- $D_t$ : Total data volume (GB or records)
- $R_p$ : Processing rate (GB/hour or records/hour)

#### 8.1. Software Solutions

New technology is increasingly being used to detect and prevent fraud before it occurs. These new proactive systems available to companies involve data mining, trend analysis, and predictive analytics that help identify vulnerabilities to fraud. Proactive detection involves testing and analyzing a company's accounting records to monitor surveillance of reviews for fraud by running software programs against an entire set of accounting data. One example of software that helps auditors analyze and visualize accounting data through Interactive Data Exploration, scripts, and visualization of a company's fraudulent schemas and patterns. The use of this software has enabled a major aluminum and recycling company to have expanded views of frauds and led to reduced auditor time in the process, increased automation, and extended views of different fraud patterns. Statistics suspect healthcare fraud to be a significant problem, leading one health service company to continuously monitor its patients' safety and databases. Software was implemented for this continuous monitoring that demonstrated innovative ideas with client experiences. For this particular health analytic company, it is crucial to maintain massively extensive up-to-date databases that process thousands of daily data transactions. Transaction red flags and retrospective data have been monitored and assessed through the software. Clients are guaranteed to have their product safety analyzed before being recalled after widespread problem dissemination.

#### 8.2. Data Management Tools

Big data can be defined as data sets that are so large or complex that traditional data processing applications are inadequate. For many organizations, the big data growth has become exponential in recent years. The implication of big data has brought significant value reduce costs, better decision making, risk management and prevention of frauds in organizations . Cloud computing helps alleviate technology costs while creating new ways for individuals to consume goods and services. Cloud computing has greater flexibility than previous forms of technology thus making it ideal for handling big data projects efficiently. Analytics as a Service (AaaS) is an adopted form provider that uses Cloud computing concept. Big data is now considered a massive amount of data that arrives in real-time or near real-time and that is widely used. Big data has become a critical factor in organizations and is widely used in areas such as preventing fraud. Data management tools are defined here as tools that are able to store, process, query/remove, clean and visualize moderate and large data. ETL tools are data management tools that can acquire and transform data from several stores to a data warehouse for further queries/reports. Integration of those ETL tools with big analytics would allow organizations easier access to tools.

When a cost trade-off is performed, Pentaho is favored in most of the situations and Sisense should be employed when an organization needs significant scale. Currently, data warehouses are one stop shops where most of the data is computed for reporting purposes. Then a report tool is used to visualize the data from the warehouse. But currently cloud services can also do a larger part of the computational role making it readily available for organizations that do not have the processing power to keep a big data. A report tool can be used to visualize data on the cloud and in addition some have the capacity to work on real time monitoring data. Thus a tool like Tableau is favored in case data is also visualized on the cloud or interest in monitoring data. Several tools can act on big data but there are restrictions on portability and integration which require increasing usage. In addition big data stores themselves can be hard to understand which means current tools should adapt and grow with the average user in mind or big data would not be unanimously used even when some relatively small organizations use it.



## 9. DATA PRIVACY AND SECURITY CONCERNS

The State of Michigan and public agencies across the nation have begun to recognize, albeit slowly, the importance of analyzing data. In particular, both predictive analytics and big data architecture provide a different, and sometimes better, way of analyzing information to deliver actionable intelligence to the end-users. Already these entities are further investing in the establishment of IT infrastructure required for analytics. However, the analysis of the availability of IT infrastructure and software vendors is of little concern if the data is not privacy compliant or secure. In fact, as sophisticated and innovative the analytics solution is, it cannot resolve issues if the data is sensitive, private, and outdated.

Privacy is a major issue in the field of predictive analytics. The field of machine learning (ML) has been continuously evolving and has yielded great success on standard corpora of data like speech recognition, image classification, and addressing synonymous phrases in search engines. However, applying the same techniques to financial management and fraud detection with government data presents unique challenges since government data is often messy, noisy, and, most importantly, privacy sensitive. Research has shown that even state and federal tax returns, which are strictly protected by law, can leak sensitive information. There is a delicate balance between applying analytics to sensitive, but potentially helpful, data. Some advanced solutions have been developed to enable privacy-preserving machine learning such as differential privacy and Multi-Party Computation (MPC). With the rise of predictive analytics and ML, the target group of the agencies needs to be re-trained to harness results effectively.

Concerns regarding cyber-attacks and breaches have emerged during government analytics initiatives. Cybersecurity will become a major and necessary focal point in the successful and effective use of big data solutions. As advanced as vendor solutions are, they are irrelevant if not deployed on secure IT infrastructures. Numerous vehicular attacks and breach incidents have occurred in public agencies in recent years even though the agencies were advised about them in advance. Security-awareness training for key personnel is required to limit the vulnerabilities and damages of the organizational breaches. Research must also be committed to further exploring innovative cybersecurity improvements.

### 9.1. Regulatory Compliance

In Australia, regulatory compliance is crucial for government agencies at both State and Federal levels. Government agencies must not only comply with the law but also have the right audit trails in place to verify compliance and attestation processes. Compliance triggers a risk-based approach to auditing. Agencies with extensive regulatory compliance requirements and controls are subject to strict statutory compliance audits, while for those with very few requirements or controls, reporting is less stringent. Compliance efforts must be proportionate to the risk of adverse outcomes and the degree of regulatory control to which the agency is subject. Achieving compliance with regulatory obligations involves identifying and gathering related evidence across the agency. The amount of evidence required, the types of continuous monitoring and reporting necessary to validate that evidence, and the attestation process are highly variable depending on the compliance regime. Some regimes strictly mandate regular attestation, while others are more adaptive to changing risks and the operating environment. Agencies must respond to inquiries, testing, and evidence requests from independent auditors engaged by the regulator to attest to compliance. Responses can take multiple forms, including formal reporting, testing and assurance of controls, inspections, and documentary evidence. The subsequent scrutiny of evidence is contentious. Regulators assess evidence using either qualitative or quantitative measures, such as prioritizing evidence types, defining tolerances against the quantity of evidence submitted, and grading evidence characteristics.

Australia's intellectual property is its enabling regulation, the PGPA Act. The PGPA Act provides responsibility for financial compliance, requiring an outcomes-based understanding of financial risk rather than prescriptive controls or accounting obligations. It remains agnostic towards continuous evidence-gathering and reporting. The PGPA Act effectively gives agencies two bodies of regulation to comply with: what they attend to in terms of their enabling legislation (the PGPA Act itself) and what evidence they must provide regarding their compliance with their legislation (the gazetted PGPA Rule, and especially the compliance checklist schedule to it). Some controls that take place in compliance processes take place in a manner prescribed heavily by the gazetted PGPA Rule for agencies with greater compliance requirements. Some agencies only have to submit accounts and the annual certificate of compliance without signifying what their reports will include.

### 9.2. Best Practices

Governments across the United States and other countries have begun to adopt enterprise resource planning (ERP) systems. While these systems can produce better information for decision-making, detect anomalies in data resulting from keying in mistakes or errors in algorithmic computations, they do not detect fraud, corruption, or abuse. This paper studies how governments can more effectively track and manage \$4 trillion in workflows for avoiding such fraud through





automatic vendor fraud detection systems built on analytic processing and data warehousing technologies. It describes a government department's system design and its capabilities that have been used in both preventing fraudulent payments and proactively identifying potentially fraudulent vendors. By using case studies, the paper also illustrates the big data infrastructure needed for such systems and issues encountered in design and implementation. The paper concludes by discussing how big data infrastructures can evolve for detecting other types of fraud.

Cash frauds are well explained as they consist in obtaining money or its equivalent under false pretences, in the form of payments never made. The vendor fraud in the computer age is in some ways similar to cash frauds. Suppose before computers were invented that a transaction recorder in a state department office purposely set up multi-statewide vendors under fictitious names with bank accounts opened as far as possible from its office. Through mails and telephones, it would receive inaccurate award orders from procurers, prepay the required supplies, and arranged for no deliveries. The vendor frauds feeding its corruptive lifestyle would thereby be quite difficult to expose by regular independent audits. In the computer age, a similar evil can be achieved as long as the auditors do not directly touch the recorders' known transactions.

## 10. FUTURE TRENDS IN GOVERNMENT FINANCIAL MANAGEMENT

Government financial management has evolved over the past few decades. The main changes are increased automation of government processes and increased use of the Internet for processing and sharing information, such as e-Government processes. When automation first started in government organizations, each automation related to a particular business process was isolated and done with information processing technology that was relevant to that business process. For example, in case of headquarter organizations, generally, a different information processing technology was applied for automation of accounts payable, cash management, asset management, and others. Gradually views started changing, particularly in the banking industry, financial organizations, and other business organizations, and IT infrastructures, software, and data warehouses started integrating.

Most advanced business organizations started concentrating on intelligent exploitation of all aggregated data for performance enhancement of all business processes and fraud detection. Government organizations still lag in integrating and intelligent exploitation of all aggregated data and process level performance enhancement and fraud detection. As there is overall interdependence across all processes in government organizations which generally handle public finance due to which fraud detection analysis needs to be aggregated process level and intelligence based. Machine learning, predictive analytics/domain intelligence, and network-enforced IT governance techniques need to be applied in a complementary way for integrating and intelligent exploitation of all data for performance enhancement of overall government financial management processes and acquisition of fraud detection analytics.

As all government processes need to be under overall structured big data-enforced IT governance. Hence the overall planning, transactional, and analytics IT infrastructure of government financial management should be structured, and all processes should abide by the outlined governance framework for greatly reducing the scope of fraud occurrence and detection. Hence both overall IT infrastructure and predictive analytics fraud detection modeling need to be properly planned and configured. Until presently, development of advanced cash management, revenue, and, to some extent, accounting process level fraud prediction modeling is desired by government organizations.

### 10.1. Emerging Technologies

Advancements in technology has offered growing capabilities for organizations to analyze and leverage vast amounts of data, resulting in improved performance and competitive advantage. Organizations are increasingly using predictive analytics to anticipate events and act proactively to improve business performance ranging from customer targeting in marketing, risk management in banking, supply chain design and retailing, product pricing in e-commerce, as well as performance management in financial management and fraud detection. Government agencies ranging from city and county governments to federal ones have also been increasingly investing in data-driven business intelligence and decision-making systems for performance management and fraud detection capabilities. No longer could agencies just rely on hindsight past reports to understand key performance indicators, trends, and expectations, but needed to search for better tools to analyze complex data landscape to monitor spending and prevent fraud using predictive analytics. Predictive analytics refers to the capability of uncovering patterns from vast amounts of historical or real-time data using statistical modeling technologies, generating predictions for better business and decision outcomes.

In recent years, the explosive growth of big data has presented new challenges for organizations to process and analyze complex heterogeneous data from various internal and external partners, leading to the growing adoption of big data technologies for tool and platform capabilities. Furthermore, the growing popularity of social media provides new streams



of unstructured data for organizations to better understand the emotions and intents of partners at a large scale that were previously more difficult and cost-prohibitive to monitor. Social network analysis has thus emerged as a new research domain and business intelligence applications area, and better tool capabilities, analytic methodologies, and control patterns have been developed as research and platforms in this domain. Despite the prior efforts to study business analytics for performance management and audit analytics for fraud detection capabilities, little research has been done to understand how to better leverage emerging technologies such as big data, text, and social media analysis technologies to improve financial management and fraud detection capabilities.

### 10.2. Predictions for the Next Decade

With the advancement of huge data generation and data handling capability, Machine Learning and Probabilistic modelling enables an immense opportunity to employ predictive analytics platform in high security critical industries namely data centers, electricity grids, utilities, airport etc. This paper proposes a novel, complete architecture of an intelligent predictive analytics platform, Fault Engine, for huge device network connected with electrical/information flow. Three unique modules, here proposed, seamlessly integrate with available technology stack of data handling and connect with middleware to produce online intelligent prediction in critical failure scenarios. The Markov Failure module predicts the severity of a failure along with survival probability of a device at any given instances. The Root Cause Analysis model indicates probable devices as potential root cause employing Bayesian probability assignment and topological sort. Finally, a community detection algorithm produces correlated clusters of device in terms of failure probability which will further narrow down the search space of finding route cause. The whole Engine has been tested with different sizes of network with simulated failure environments and shows its potential to be scalable in real-time implementation [7].

Taken together, these studies provide remarkable support for the validity of intelligent financial fraud detection in terms of recent developments. However, despite this, there is still a tremendous gap to be filled in a number of areas preliminary to formulating an integrated template of best practice. Financial fraud detection is an evolving field in which it is desirable to stay ahead of the perpetrators. Additionally, it is evident that there are still facets of intelligent fraud detection that have not been investigated. Typical classification problems CI and data mining-based financial fraud detection is subject to the same issues as other classification problems, such as feature selection, parameter tuning, and analysis of the problem domain. Information misrepresentation applies to several types of financial fraud classification problems, including financial statement fraud, revenue recognition fraud, and earnings management fraud. Detection of fraudulent reports and trailing events respective efforts have been made to expose financial fraud via the detection of manipulated financial statements and network events .

## 11. STAKEHOLDER ENGAGEMENT

Facilitating effective stakeholder engagement when integrating predictive analytics with IT infrastructure is crucial to gaining the support of high-ranking executives and end users. This requires addressing major points of concern that figure prominently in the minds of stakeholders, as well as some points that are not so obvious. For example, stakeholders should be assured both that the proposed technology is applicable to the relevant context and that it is feasible to implement . Regarding feasibility, stakeholders should be reassured that the proposed approach does not necessitate impractically large-scale changes to existing systems and processes. When these reassurances are credible, it is easier to persuade stakeholders of the expected benefits of the offered solutions. Although the latter is considered essential for gaining stakeholder support, it is also the aspect that practitioners say is difficult to assess, especially in the public sector context.

A promising way to assess expected benefits of stakeholder engagement and participation is using an evaluative framework of key stakeholder engagement parameters. This allows tailoring the approach to the specific goals and circumstances of the engagement. Existing stakeholder engagement studies do refer to relevant parameters. One way the proposed approach can inform scholars and practitioners regarding effective stakeholder engagement is to develop these parameters semi-independently from a concrete stakeholder engagement implementation. Instead, it can be considered a systematic approach that provides a useful framework that consolidates existing insights regarding benefits of anticipated participation, facilitating integration with the specific stakeholder engagement context by adding context-specific solutions. While it is hardly possible to provide a fully systematic analysis of appropriate parameters here, a structured initial inventory is considered valuable.

The proposed evaluation framework can be used to identify points of concern among the stakeholders. This helps identify which concerns should be acted upon in a concrete implementation of the engagement. It can also serve as a basis for tailored actions. For example, in the case of voluntary engagement of decision makers, emphasising applicability and



feasibility may make it feasible to present expected benefits convincingly. Alternatively, in the case of a more mandated type of engagement, emphasising expected benefits can still be valuable in framing the recommendations proposed to the stakeholders engaged initially.

### 11.1. Identifying Key Stakeholders

In the migration of predictive analytics (PA) installed solutions to the cloud, not only does the availability of the data and underlying system architecture change, but also the responsibility and ownership of the critical data required for further development of the models. Furthermore, the system architecture paradigm itself changes from a classical on-premise hosted, system closely managed and supervised by internal IT personnel to a cloud vendor controlled system architecture. Therefore, the information technology (IT) setting resulting from the migration of an advanced output modeled system to a cloud solution bears a risk in which critical knowledge and decision authority might be lost. Thus, this requires and provides the first condition within the framework, the identification of key stakeholders. Five dimensions for identifying key stakeholders are proposed .

First and foremost, the structure dimension identifies stakeholders involved in the architecture of the PA models and supporting data and system environment. Secondly, the decision dimension identifies stakeholders with decision power regarding further use, development, enhancement, or modification of the data, models, or tools in the IT infrastructure. Subsequently, the usage dimension identifies stakeholders using the generated outcomes of the models, supporting tools, or underlying data. Furthermore, the service dimension identifies stakeholders providing services or information regarding the data, models, or tools. Finally, the concern dimension identifies stakeholders who hold a personal interest (positive or negative) regarding the PA intelligence or supporting tools, models, or rationale. The first aspect for each dimension of every stakeholder is drafted separately to provide an unbiased overview of stakeholders within the domain of study.

### 11.2. Strategies for Engagement

The World Bank has implemented a set of six mandates to tackle corruption in banks, and in 2000, the development of a new financial management information system was initiated by launching the "Greening Public Financial Management" (PFM) program strategy, which emphasized advanced and innovative IT-enabled PFM reforms. Corruption is a root cause of poverty in the World Bank's internal documents. With increasing decades of experience, using the lessons learnt over the years and taking possible actions, effective strategies for new engagement in financial management PFM systems upgrading in governments are recommended hereafter.

Establishing a formal anticorruption steering group executive task force is the first strategy recommended to engagement in government PFM systems upgrading. A multi-disciplinary team is needed to drive the anticorruption policy and specific measures for their implementation. The steering group should establish a fraud recovery litigation committee and recover the fraudulently taken amounts as a first priority afterwards. Some measures on recovering stolen assets and addressing employees' misconduct should be taken immediately. Establishing a steering group helps take an overall perspective of anticorruption, allowing the anticorruption initiatives to become part of the policy agenda framed in terms of government priorities. After conducting an in-depth fraud situation assessment and other scheduling actions arranged, the steering group can turn its focus to the high-level strategizing of anticorruption, especially the recommendation of a package of new anti corruption measures.

A detailed anticorruption roadmap and a package of anticorruption measures are the next strategy. In recovering initial asset lost to fraud, coordination and communication of executive decision, inter-agency cooperation, government department collaboration, and partnership with other organizations are important and challenging. A package of adjusted anticorruption measures should be gradually revealed to the public in proper times in anticipation of another wave of protests. Such measures include the development of IT-led and data-driven analysis corruption prevention measures; establishment of anticorruption institutions outside of government; adjustment of anti-corruption government institutions' authorities; encouragement of journalists to investigate the exposed corruption; and international collaboration in investigating corruption and recovering illegal assets [4].

## 12. TRAINING AND CAPACITY BUILDING

As budget reform and digitalization in public finances are often initially seen as purely technical issues, when viewed from an information technology perspective, the best tools and systems enabling reform are mostly available. However, IT systems are merely enablers; managing reform successfully depends on individuals and organizations. Thus, the focus



must be on wider management issues, rather than purely technical ones. In addition, government financial management reform often means a paradigm shift. Such profound changes, in turn, require not only new processes and systems but also that individuals and organizations adopt new views and thinking in line with the new paradigm and become capable of executing and utilizing the new systems. Efforts at individual and organizational levels are often required, and time may be needed for successful change. The success of efforts often depends on timing and experience. Therefore, capacity building in terms of training and other measures must be a central part of the implementation approach. It is important to recognize that wider management change will mostly utilize new existing tools and these should be carefully selected. Attention must be given to the content of training, e.g. if not simply focused on the buttons to be pressed on the newly established systems. It should be remembered that in the longer run the systems will change, new versions will be implemented, and operations will become even more complex. As a result, knowledge on substantive issues, underlying principles, budget reform ideas, and new ways of thinking must be included in training. Training should focus little on how button pressing is done and mostly on new issues like forecasting and accruals, which will not form part of training provided by vendors. The reform, which encompassed almost everything, was divided into packages. The central on-line cash management system was delivered as planned, and implementation was fairly problem-free. Six-month income and expenditure forecasting routines were developed in relatively good time, challenging the old attention to historical statements. New monetary rules were brought in, nation-wide practices developed, and ongoing communication was established between revenue-collecting institutions and the Ministry of Finance. Methods on accruals were also developed in many places with broadly similar content.

### 12.1. Skill Development Programs

Skill development programs envisaging cross-organization and cross-national participation need to be initiated to deal with the challenges and opportunities in research with big data. Some innovative strategies are necessary for forecasting the behavior of decision-making systems or complex networks in which many interactive elements are operating together and simultaneously. A fuzzy cognitive map (FCM) is a causal-oriented graph used as a knowledge representation tool to model and simulate the high-level behavior of an interactive decision system. Moreover, an FCM has numerous applications ranging from decision support and process simulation to strategic marketing planning in industrial firms. An FCM allows any user to analyze complex systems of mostly qualitative nature.

An FCM can serve both predictive and diagnostic purposes. In the predictive application, each node may expect its value to forecast future outcomes. Because of the recursive architecture, the prediction should be undertaken repeatedly. With this learning capability, the FCM can be considered a type of neural network, and it would be interesting to do analytical work upon it. Nevertheless, as it can be properly used on qualitative issues, the FCM can also provide a major advantage over neural networks, which can assure good performance but do not need to offer explanations for their outputs. This aspect is crucial in decision-making situations where pathological behavior would severely worsen the consequences, as in stock forecasting. Here, the why is at least as important as the what. Generally accepted reasons for occurrence are always preferred over sheer predictions even when very precise. This way, the FCM can use human expertise but provide more extensive accounts on modeling and decision issues than any human expert could provide.

In the diagnostic application, the model produces endogenous values from fixed exogenous ones. Here, the distinction of being a time series is irrelevant. The predefined structure is fixed, and the model will behave accurately in real-time as long as the input remains within bounds specified by the model derivation. In such a way, both short-term detection tasks and system monitoring decisions can be addressed. Typically in monitoring tasks, the static state of the system will affect an outcome, while a certain input will have an impact on the output only with a time delay. In procurement decisions as in many socio-economic applications, the input-output relationships of the overall process tend to be static in the long run, while exogenous variables continuously fluctuate.

### 12.2. Organizational Change Management

The approach taken to assist and understand the changes that occurred as a result of the original Concert implementation and subsequent finish of the implementation and replacement with the new Dance system pertained to employees' perceptions of those changes. This included what initiatives were successful, what initiatives were deemed to be poor, what unintended consequences arose, and what employee perceptions led to these outcomes. It was broadly acknowledged that change management was not a success and an apparent disconnect between what management believed occurred and what employees perceived to have occurred. As a starting point there was a broad understanding that IS adoption and organizational change are related and interconnected efforts that are a concern to IS researchers and practitioners. Through discussions it was agreed that what occurred in regard to change management of Concert could not and can not be easily seen in isolation from what occurred in regard to change management initiatives taken with respect to Dance.





An understanding of both adoption from the perspective of the individual and organizational change management or planned organizational change is therefore needed. As a result it was decided as a first step it was first necessary to understand what changes had occurred as a result of the implementation of the new Dance system. Previous attempts at organizational change management were seen to lack success. Management could not provide a list of initiatives or the individual changes that were intended to occur [14]. A dataset that might assist in understanding and planning was therefore not available. Without an understanding of what changes were intended and what actually occurred there was little hope of effectively understanding or re-planning change initiatives for the ongoing projects in the Department of Internal Revenue. It was therefore necessary to conduct a survey of employees perceptions of which changes had occurred, the management initiatives that were perceived to be successful, some that were not seen to be a success and some perceived unintended consequences of the change.

### 13. EVALUATION METRICS

Financial Integrity Systems have to monitor transactions and decide whether to approve, decline, or initiate manual review. Each decision according to a list of precise business rules is often a combination of several run-time and pre-processing curves such as weighting against global activity features, weighting against historical activity, and similar curves. The rules' average inclination is determined by senior business personnel and reflects the governmental and commercial requirements. Compliance with these rules is crucial for ongoing financial activity. Therefore holding a high level of precision is important; high false positive rates mean blocking an average bad user and adding handling costs. However, this also means potential service denial and swollen chargebacks claims .

For fraud detection purposes, one must hold a significant level of recall against fraudulent users as well. The blocked fraudulent users are the most conservative estimate of the capability of the fraud detection, this pool (detection control group) ideally should be stored for a long time period. Reasons for this that inhibit long-term storage include resource cost of storage and legal considerations of users' data storage and retention period to comply with common laws and regulations worldwide.

Modern Financial Integrity System's engine used for fraud detection is typically an ensemble of algorithms including decision trees, logistic regression models, behavioral rules, and any other hosts of techniques commonly used for fraud detection in other domains. Each engine may leave out its run-time and online pre-processing curves, types of activity features, curving methods, and pre-packaged libraries, but typically consists of the basic algorithms above. Training such a predictive model may take hours or days, therefore developing a new forensic algorithm can hardly be seriously evaluated in online A/B tests.

#### 13.1. Key Performance Indicators

Finance departments in Government organizations can be challenged if a high-volume, structured, data warehouse database does not exist. Also called a Data Lake, this repository will store all types of structured, semi-structured, and unstructured data in its source format. It will allow different types of analytic workloads and the usage of AI/ML applications over a vast database, offering new insights that have not been explored before.

If such data repositories already exist, it makes sense to leverage them, letting knowledge management protocols take care of the configuration, naming, and structure of the table. It is not to say that the usage of a data warehouse is a goal in itself. It does make more sense to stake all the available data in its original format, but if an organization feels comfortable with it, a Data Warehouse is a valid option. So perhaps at first most of the construction processes have best practices and adjustment options based on available tables. If high-level KPIs already exist, the construction has to insert data in a proper "region" or layer, run ETLs, and store them in a Data Warehouse.

Given the nature of public organizations and the Finance departments, there are always limitations in terms of public knowledge and comparative data with other organizations. Since computing storage is not an issue in these times, reverse-engineered Linear Regression models would offer a wide range of established KPIs to place all the records.

A Monitoring Dashboard focusing on Government organizations requires the customization of customary designs and requirements. However, all its panels will be based on core KPIs already considered good practices in Finance departments – in Government organizations specifically, banks, and acquisition houses [15]. Basic price distribution curves for expenses and investments would need little customization but would offer immediate insight into financial behavior and controls.





### 13.2. Measuring Success

When a data analytics project is proposed, the management receiving the request is often interested in how the technology would impact the business. Accounting departments in particular focus on its control over costs and resources. It is widely accepted that data analytics impacts quality and understanding of the analyzed data. However, measuring this business impact is difficult. One method is to analyze the impact of expectations on the projects' success. With this, some common patterns were found in the data analytics projects considered, which are correlated with differences in their outcomes. In analyzing the cases based on objective criteria, two constructs were derived, which relate to the planners' ambitions and the projects' scopes. Projects with exceedingly high ambitions failed because they did not sufficiently focus on making data analytics operational. Apart from purely practical-driven projects, two other strategies were found that could succeed in either focusing on objectives aligned with the goals of controlling or of obtaining new information. Examining this allows a better understanding of the mechanisms underlying the projects' outcomes and improves knowledge on data analytics projects' success in turn. This work centers on the successful implementation of data analytics projects in companies. The resulting understanding helps answer the question of which broad strategy is most likely to lead to (positive) project outcomes and whether strategy had a significant impact on project results. The sample consists of data analytics projects across a variety of industries in public and private sectors and with diverse analytical techniques and outcomes [16].

## 14. POLICY RECOMMENDATIONS

Integrating Predictive Analytics and IT Infrastructure for Advanced Government Financial Management and Fraud Detection

### 14. Policy Recommendations

Public organizations and governments hold a lot of valuable data that could be exploited for wide-ranging advantages. Although in recent years most entities of the public sector have taken steps to develop their IT infrastructure, the exploitation of analytics goes far beyond that. High-value databases have to be linked and processed by a data science team to develop useful deliverables. All major departments of both national and local governments could gain a lot by enriching their current systems with a tactical business intelligence tool. This tool would help administration staff to monitor and analyze KPIs and take timely decisions to optimize missions, compliance, and financial aspects. Those KPIs would stem from various databases and could be visualized through easy-to-understand dashboards.

Since fraudsters are very creative, it is hard to find a universal detection mechanism based on key indicators. Nevertheless, public organizations could be equipped with a powerful screening environment that would be able to issue red flags by predicting suspicious situations or behavior. By leveraging the entity's historical databases along with scalable state-of-the-art ML algorithms, such environments could be developed as a cloud-based platform with a proven case. The development of a machine-learning cloud environment for predicting fraudulent loans and bridging that with public-sector financial management would be a very ambitious and challenging task with a big potential. While several application domains are exploiting the added-value of analytics over various datasets to obtain actionable insights, the public policy management domain has not yet taken advantage of the full potential of analytics and data models to realize efficient policy management [1].

#### 14.1. Strategic Policy Framework

Currently, dynamic environments characterized by complexity, uncertainty, and highly important and interconnected issues prevent the straightforward development and evaluation of public policy decisions. There is an increasing necessity for prediction/counterfactual analysis capabilities that go beyond more straightforward response or scenario analysis. Several application domains, such as urban planning, transport, environmental, and public policy management, have been exploiting the data-driven modeling and analytics methods and techniques. It is an overall architecture of a cloud-based environment enabling analytics as a service, focusing on the facilitation of data-driven public policy management [1]. Data collection from heterogeneous sources, linking, and aggregation are provided, complemented with data cleaning and interoperability techniques.

Moreover, interactive visualization and exploratory data analytics tools are offered, enabling comprehensive results' inspection. An innovative framework implementing different predictive and descriptive modeling methods is presented, fulfilling the user's requirements for accurate, explainable, and impartial results. Finally, an evolving conceptual framework of a Policy Development Toolkit is presented, linking the analytics as a service cloud computing environment with forethought policy development, creation, and optimization tools. These tools are being developed based on state-



of-art academic and practical methods and constitute first attempts to exploit advanced analytics and counterfactual analysis capabilities in retail/fine grain transport planning policy analysis and public policy development. At the first level, the delivered analytics as a service cloud computing architecture will create a data and policy integrated environment enabling the exploitation of adequate databases to develop analytics solutions fulfilling public policy management purposes. At the second level, the development of innovative analytics methods and tools complying with the formed general architecture will be performed.

#### 14.2. Implementation Guidelines

The implementation process of such predictive analytics and IT infrastructure projects in the context of government financial management is inherently complex and can encompass a variety of tasks. It starts with problems/signal identification, where organisations review available data on past, current and planned events including accounting issues, legislation issues and financial portfolio issues. The second step is data collection, including the harvesting of data from previously mentioned events and data stored in archives. Due to large unstructured data volumes, a specific approach or instrument is needed to maintain quality and effectiveness, i.e., natural language processing techniques on texts of huge size or sound-wave analysis software on speeches of respective officials. The third stage is data transformation. Raw data gathered in the second step often needs a thorough upgrade in order to address the previous data insufficiency issues with completeness, consistency, time, uniqueness and validity. These will require advanced data wrangling processes to employ statistical or numerical methods in tedious cleaning, aggregating, calling, regularising or matching tasks. Subsequently, the software development phase results in the visualisation and user interface design for the predictive analytics process to be easily exploited by end users without artificial intelligence or statistical background [16]. The last step is intelligence improvement, which constitutes improving the extraction processes by identifying counterparts, adding queries, and integrating new technologies and frontiers such as blockchain and big data.

The evolving government character of the organisation can give stronger priority to the aspects of adopting software than controlling it, and is supposed to be of strong mutual interdependence, i.e., uncovering the shortage of predicting ability makes it possible to detect ignorance of a software solution in extracting and analysing data. The effect of the post-implementation stages is expected to be easy to access but profound or damaging, if ignored, e.g., further behavioural entitlement of adverse consequences or the neglect of identification due to software malfunctioning [2]. Expert systems, as a type of artificial intelligence software, conduct predictive analytics tasks by self-processing raw data on predefined reasoning rules. This off-the-shelf software type requires no coding while the domain model built into systems largely limits the modelling of scenarios. Nevertheless, this risk can be lowered by searching for software with a domain model that can be modified or prepared in advance with reliable knowledge by experts in the respective jurisdiction. The scoping step of signal identification, where legal knowledge is needed to extract data smartly from text documents, is a crucial user interaction process in this integration type. Complexity Preview such controlling action can be rendered automatic by using no-amended off-the-shelf software and thus be out of reach by the user, but the consequence is high devastation of trust (e.g., misuse of authority & grave mistakes).

#### 15. CONCLUSION

The combination of predictive analytics and IT infrastructure provides advanced state-of-the-art tools that public sector governments can use to support financial management and detect fraud. Predictive analytics encompasses a set of tools and techniques that allow decision-makers to analyze past data in making better decisions for the future. Thus, it can be leveraged to uncover trends and patterns in historical data and then project what will happen next. On the other hand, an intelligent IT infrastructure including advanced technology such as cloud computing, big data, data warehousing, and data lakes enables government finance departments to receive, integrate, and utilize enormous volumes of big data across multiple departments and jurisdictions. The convergence between predictive analytics and IT infrastructure will facilitate massive, holistic, and real-time modeling and simulation of the current states, trends, and patterns of the various financial management processes and performance outcomes. Extending a conventional financial data warehouse into a data lake integrates, stores, and preprocesses vast financial data generated or collected from various sources at a granular level. State-of-the-art data processing engines automate the computation of the financial statistics that feed the predictive analytics engine. All these new technology enablers will modernize public sector government financial management, which will shift from focusing on reporting and auditing past historical data and compliance measures to proactively managing performance and detecting any process faults, errors, and frauds that may occur in the present or future.

Moreover, a predictive analytics engine will issue alerts about potential dangerous scenarios so that government financial managers may act and intervene as soon as possible. Hence, financial management and fraud detection will no longer be a “behind-the-scene” event; rather, they will have boardroom significance and become an integral part of the ongoing



operations like budgeting, accounting, and reporting. Such holistic improvements will reform the business processes of public sector government finance departments and hence dramatically improve their overall efficiency, effectiveness, transparency, accountability, integrity, and credibility.

## REFERENCES

- [1] Chava, K., Chakilam, C., Suura, S. R., & Recharla, M. (2021). Advancing Healthcare Innovation in 2021: Integrating AI, Digital Health Technologies, and Precision Medicine for Improved Patient Outcomes. *Global Journal of Medical Case Reports*, 1(1), 29–41. Retrieved from <https://www.scipublications.com/journal/index.php/gjmcr/article/view/1294>
- [2] Nuka, S. T., Annareddy, V. N., Koppolu, H. K. R., & Kannan, S. (2021). Advancements in Smart Medical and Industrial Devices: Enhancing Efficiency and Connectivity with High-Speed Telecom Networks. *Open Journal of Medical Sciences*, 1(1), 55–72. Retrieved from <https://www.scipublications.com/journal/index.php/ojms/article/view/1295>
- [3] Avinash Pamisetty. (2021). A comparative study of cloud platforms for scalable infrastructure in food distribution supply chains. *Journal of International Crisis and Risk Communication Research*, 68–86. Retrieved from <https://jicrcr.com/index.php/jicrcr/article/view/2980>
- [4] Anil Lokesh Gadi. (2021). The Future of Automotive Mobility: Integrating Cloud-Based Connected Services for Sustainable and Autonomous Transportation. *International Journal on Recent and Innovation Trends in Computing and Communication*, 9(12), 179–187. Retrieved from <https://ijritcc.org/index.php/ijritcc/article/view/11557>
- [5] Balaji Adusupalli. (2021). Multi-Agent Advisory Networks: Redefining Insurance Consulting with Collaborative Agentic AI Systems. *Journal of International Crisis and Risk Communication Research*, 45–67. Retrieved from <https://jicrcr.com/index.php/jicrcr/article/view/2969>
- [6] Singireddy, J., Dodda, A., Burugulla, J. K. R., Paleti, S., & Challa, K. (2021). Innovative Financial Technologies: Strengthening Compliance, Secure Transactions, and Intelligent Advisory Systems Through AI-Driven Automation and Scalable Data Architectures. *Universal Journal of Finance and Economics*, 1(1), 123–143. Retrieved from <https://www.scipublications.com/journal/index.php/ujfe/article/view/1298>
- [7] Adusupalli, B., Singireddy, S., Sriram, H. K., Kaulwar, P. K., & Malempati, M. (2021). Revolutionizing Risk Assessment and Financial Ecosystems with Smart Automation, Secure Digital Solutions, and Advanced Analytical Frameworks. *Universal Journal of Finance and Economics*, 1(1), 101–122. Retrieved from <https://www.scipublications.com/journal/index.php/ujfe/article/view/1297>
- [8] Gadi, A. L., Kannan, S., Nandan, B. P., Komaragiri, V. B., & Singireddy, S. (2021). Advanced Computational Technologies in Vehicle Production, Digital Connectivity, and Sustainable Transportation: Innovations in Intelligent Systems, Eco-Friendly Manufacturing, and Financial Optimization. *Universal Journal of Finance and Economics*, 1(1), 87–100. Retrieved from <https://www.scipublications.com/journal/index.php/ujfe/article/view/1296>
- [9] Cloud Native Architecture for Scalable Fintech Applications with Real Time Payments. (2021). *International Journal of Engineering and Computer Science*, 10(12), 25501-25515. <https://doi.org/10.18535/ijecs.v10i12.4654>
- [10] Pallav Kumar Kaulwar. (2021). From Code to Counsel: Deep Learning and Data Engineering Synergy for Intelligent Tax Strategy Generation. *Journal of International Crisis and Risk Communication Research*, 1–20. Retrieved from <https://jicrcr.com/index.php/jicrcr/article/view/2967>
- [11] Chinta, P. C. R., & Katnapally, N. (2021). Neural Network-Based Risk Assessment for Cybersecurity in Big Data-Oriented ERP Infrastructures. *Neural Network-Based Risk Assessment for Cybersecurity in Big Data-Oriented ERP Infrastructures*.
- [12] Katnapally, N., Chinta, P. C. R., Routhu, K. K., Velaga, V., Bodepudi, V., & Karaka, L. M. (2021). Leveraging Big Data Analytics and Machine Learning Techniques for Sentiment Analysis of Amazon Product Reviews in Business Insights. *American Journal of Computing and Engineering*, 4(2), 35-51.
- [13] Routhu, K., Bodepudi, V., Jha, K. M., & Chinta, P. C. R. (2020). A Deep Learning Architectures for Enhancing Cyber Security Protocols in Big Data Integrated ERP Systems. Available at SSRN 5102662.
- [14] Chinta, P. C. R., & Karaka, L. M. (2020). AGENTIC AI AND REINFORCEMENT LEARNING: TOWARDS MORE AUTONOMOUS AND ADAPTIVE AI SYSTEMS.