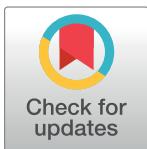


RESEARCH ARTICLE

Bayesian inference of local government audit outcomes

Wilson Tsakane Mongwe^{1,3*}, Rendani Mbuva^{2,3}, Tshilidzi Marwala^{1,3}

1 School of Electrical Engineering, University of Johannesburg, Auckland Park, South Africa, **2** School of Statistics and Actuarial Science, University of Witwatersrand, Johannesburg, South Africa, **3** Johannesburg Machine Intelligence Lab, Johannesburg, South Africa

* wilsonmongwe@gmail.com

Abstract

The scandals in publicly listed companies have highlighted the large losses that can result from financial statement fraud and weak corporate governance. Machine learning techniques have been applied to automatically detect financial statement fraud with great success. This work presents the first application of a Bayesian inference approach to the problem of predicting the audit outcomes of financial statements of local government entities using financial ratios. Bayesian logistic regression (BLR) with automatic relevance determination (BLR-ARD) is applied to predict audit outcomes. The benefit of using BLR-ARD, instead of BLR without ARD, is that it allows one to automatically determine which input features are the most relevant for the task at hand, which is a critical aspect to consider when designing decision support systems. This work presents the first implementation of BLR-ARD trained with Separable Shadow Hamiltonian Hybrid Monte Carlo, No-U-Turn sampler, Metropolis Adjusted Langevin Algorithm and Metropolis-Hasting algorithms. Unlike the Gibbs sampling procedure that is typically employed in sampling from ARD models, in this work we jointly sample the parameters and the hyperparameters by putting a log normal prior on the hyperparameters. The analysis also shows that the repairs and maintenance as a percentage of total assets ratio, current ratio, debt to total operating revenue, net operating surplus margin and capital cost to total operating expenditure ratio are the important features when predicting local government audit outcomes using financial ratios. These results could be of use for auditors as focusing on these ratios could potentially speed up the detection of fraudulent behaviour in municipal entities, and improve the speed and quality of the overall audit.

OPEN ACCESS

Citation: Mongwe WT, Mbuva R, Marwala T (2021) Bayesian inference of local government audit outcomes. PLoS ONE 16(12): e0261245. <https://doi.org/10.1371/journal.pone.0261245>

Editor: Thippa Reddy Gadekallu, Vellore Institute of Technology, INDIA

Received: May 16, 2021

Accepted: November 25, 2021

Published: December 14, 2021

Copyright: © 2021 Mongwe et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are available on the national treasure of South Africa website: <https://municipaldata.treasury.gov.za/>. This data is already available at this URL.

Funding: Funding for this research was provided by Google PhD Fellowships and the National Research Fund of South Africa. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

1 Introduction

The Auditor General of South Africa (AG-SA) revealed that South African local government entities lost over \$2 billion in irregular expenditure in the 2018-2019 financial year [1, 2]. This has consequently had a negative impact on service delivery and returns on the rapidly increasing government debt [1]. The manipulation of financial statements is not only limited to the public sector, with the Steinhoff collapse being a prime example of management fraud in the

private sector [3]. Steinhoff is a South African retailer that lost over R200 billion in market capitalisation on the Johannesburg Stock Exchange (JSE) over a short space of time after allegations of accounting fraud [2, 3].

The recent scandal of Wirecard, and previously Enron, also indicate that financial statement fraud is not only a problem for South Africa, but the world at large [4, 5]. Wirecard is a German payment processing company, which filed for insolvency in 2020, that manipulated its financial statements by misstating its profit [4]. Enron was an American natural gas company that lost over \$60 billion in market capitalisation in the early 2000s after the allegations of fraud emerged [5].

The use of automated techniques for the analysis and detection of financial statement fraud has been on the increase in the past decade [2, 6–13]. Mongwe and Malan [2, 6] outline how artificial intelligence and other automated methods can be used to construct decision support tools for various stakeholders. For example, auditors may use the decision support tool to flag entities who are at risks of having committed financial statement fraud and reduce the turn around time of audit amongst other benefits [2, 6]. A prime example within the South African context is the AG-SA having to audit all local and provincial government entities at the end of each financial year [1, 6].

Logistic regression has been successfully used in the literature for the detection of financial statement fraud [7, 14–21]. Moepya *et al.* [21] use logistic regression in the detection of fraud in companies listed on the JSE, while Boumediene *et al.* [22] performed a similar study for entities listed in Tunisia. Logistic regression has advantages over more complicated models such as artificial neural networks in that the results are more easily interpretable by the stakeholders, which is an important consideration when building a decision support tool [2, 6, 10, 21].

In logistic regression, as with any other machine learning model, one has to decide on the input features to use. Correctly selecting the variables to use as inputs for the models is important because it can influence the performance of the models [23]. Utilising feature selection techniques can improve the models predictive performance and reduce the model complexity as fewer features would be required [2, 24]. Examples of feature selection methods used in the financial statement fraud detection literature include correlation, t-test, analysis of variance, decision trees and principal component analysis [2, 21, 22, 25–28]. In this paper we limit ourselves to only using financial ratios for the prediction of financial statement audit outcomes. Thus feature selection in this context amounts to selecting which financial ratios are the most important or relevant in the inference of financial statement audit opinions.

In this work, we present the first use of Bayesian logistic regression with automatic relevance determination (BLR-ARD) for the inference of audit outcomes. The Bayesian approach allows us to measure the uncertainty in our predictions, which gives a sense of how much confidence we have in a particular prediction. The use of automatic relevance determination allows us to automatically determine which of the input features are the most relevant, with uncertainty measures around these as well [29]. This results in the model outcomes being more interpretable, allowing stakeholders to better understand the results of the model.

The motivation behind this work is in understanding the financial performance of South African municipalities in terms of audit outcomes, and particularly the features or financial ratios that drive these audit outcomes. We approach this problem from a Bayesian perspective, which is a first-in-literature, as it provides a probabilistically principled framework for predicting and understanding the audit performance of local government entities. This framework also enables us to provide uncertainty levels in the predictions produced by the models, and further allows us to automatically identify and rank the most important financial ratios for

audit outcome modelling using prior distributions—which is an important contribution of this work.

The results of our analysis can be useful to various stakeholders, but particularly the Audit General of South Africa (AG-SA) and other auditors of local governments around the world. The results indicate which financial ratios the auditors could focus on so as to efficiently identify likely instances of financial irregular behavior and high financial risk in local government entities, and thus improve the speed and overall quality of the audit. As such audits are performed with limited resources—a framework such as the one presented in this work can be used for resource allocation on the basis of the predicted risk.

We train the BLR-ARD model parameters with Markov Chain Monte Carlo (MCMC) methods. MCMC methods have an advantage over approximate inference methods, such as variational inference and Laplace approximations, as MCMC methods are asymptotically guaranteed to converge to the true target posterior distribution [30]. In this paper, we present the first use of the Metropolis-Hasting (MH) algorithm, Metropolis Adjusted Langevin Algorithm (MALA), Separable Shadow Hamiltonian Hybrid Monte Carlo (S2HMC) and the No-U-Turn Sampler (NUTS) MCMC algorithms in the training of BLR-ARD models for inference of financial statement audit opinions.

The MH algorithm suffers from random walk behaviour [31–33]. This is due to the MH algorithm simply adding random Gaussian noise to the current state to determine the next state—with the noise being independent from the noise used to generate the current state [34]. The MH algorithm is often combined with other methods to construct hybrid Monte Carlo methods such as HMC [34–36].

The MALA improves on the MH algorithm by using first order gradient information. This reduces the random walk behaviour of the MH algorithm [36]. The MALA can be shown to be a special case of HMC with the trajectory length of one [35, 37]. As with the MALA, the HMC method uses first order gradient information of the target posterior to assist the exploration of the parameter space, but also adds an auxiliary momentum variable to the parameter space to allow it to explore different energy levels [32, 35, 38]. The HMC method has various parameters that need to be tuned, which is an impediment of the algorithm being broadly used in practice [39, 40]. The NUTS algorithm with dual averaging resolves this by adaptively setting the HMC parameters [39, 40].

The S2HMC algorithm has been shown to provide better sampling behaviour when compared to HMC [38, 41–43]. This is due to the shadow Hamiltonian in S2HMC being better conserved by the leap-frog integrator. This leads to higher acceptance rates and lower autocorrelations in the generated samples [38, 41, 44]. The main drawback of S2HMC is its high execution time, which reduces its performance on a execution time normalised effective sample size basis [38].

The empirical results in this work show that the S2HMC, NUTS and MALA algorithms are able to better explore the target posterior than the MH algorithm. This results in the S2HMC, NUTS and MALA algorithms having similar predictive performance, with the MH algorithm being significantly outperformed. The S2HMC method produces higher effective sample sizes than NUTS, indicating S2HMC's ability to better explore the target posterior distribution than NUTS. The majority of the algorithms agree on which set of features are the most relevant for modeling audit opinions. The analysis shows that the current ratio, debt total revenue and the net surplus margin are the important features when predicting local government audit outcomes using financial ratios.

Note that multiple MCMC algorithms are considered in this work so as to assess the robustness of the results produced by the Bayesian framework. The results do indeed indicate that the methods agree on which financial ratios are important, suggesting that the Bayesian

approach undertaken in this study is robust. Agreement on feature relevance from the multiple inference methods is key to avoiding spurious identification of relevant financial ratios.

The main contributions of this work are as follows:

- We present the first application of a Bayesian inference approach to the problem of predicting the audit outcomes of financial statements of local government entities using financial ratios.
- We present the first implementation of the BLR-ARD trained with the Separable Shadow Hamiltonian Hybrid Monte Carlo, No-U-Turn Sampler, Metropolis Adjusted Langevin Algorithm and Metropolis-Hasting algorithms.
- Unlike the Gibbs sampling procedure that is typically employed in Bayesian sampling from ARD models [29], in this work we jointly sample the model parameters and hyperparameters. This results in a more stable exploration of the posterior.

The remainder of this paper is structured as follows: Section 2 provides the background to financial statement fraud, Section 3 discuss the Markov Chain Monte Carlo methods used in this work, Section 4 outlines the experiments conducted and the dataset used, Section 5 presents and discusses the results of the experiments and we provide the conclusion in Section 6.

2 Overview of financial statement fraud detection

The financial statements of an entity consist of reports such as the income statement, balance sheet and cash flow statements [2, 6]. These statements are usually summarised into financial ratios, and used by different stakeholders for various purposes [2, 6]. For example, investors use the financial statements to determine if the entity is a good investment while the government would use financial statements to determine the tax payable to the state by the entity [2, 6]. These financial statement are typically summarised into financial ratios [2, 6].

Financial statement fraud, or management fraud, occurs when the financial statements of an entity are manipulated so as to make the entity appear to be in a better financial state than is actually the case, as was the case with Enron, Steinhoff and Wirecard [2, 4, 6, 45–47]. This manipulation is often perpetrated by the management of the entity, and at times with the support and knowledge of the auditors of the entity—as was the case with Enron [2, 5, 6, 47]. Examples of fraud that can be present in a financial statement of an entity include the omission of material information such as a large expense and the manipulation of the entity's profits [2, 6, 13, 45].

In this paper, we model the audit opinion expressed by the AG-SA using financial ratios created from the local government entity's financial statements. The audit opinion expressed by the AG-SA on the financial statements of South African municipalities falls broadly into the following categories [2, 6, 48]:

- *Clean or unqualified audit opinion*—The financial statements contain no material misstatements. Note that this does not necessarily mean there was no fraud.
- *Qualified audit opinion*—The financial statements contain material misstatements in specific amounts, or there is insufficient evidence to conclude that the amounts are not materially misstated.
- *Adverse audit opinion*—The financial statements contain material misstatements. This however does not necessarily mean that there was fraud present.
- *Disclaimer audit opinion*—The municipality provided insufficient evidence in the form of documentation on which to base an audit opinion.

From the above, it is clear that clean or unqualified audit opinion are preferred to the other audit opinions. For the purposes of this study, we consider the statements of a municipality to be a fraudulent instance if the audit opinion is not a clean or unqualified audit, which is consistent with other studies in the literature [2, 6, 12, 49]. Thus we consider a financial statement to be fraudulent if the AG-SA expressed a qualified, adverse or disclaimer audit opinion, and a financial statement is considered not fraudulent when it receives a clean or an unqualified audit opinion [2, 6].

3 Methodology

3.1 The model

In this work, we model the local government audit outcomes using Bayesian logistic regression. The negative log-likelihood $l(D|\mathbf{w})$ function associated with logistic regression is given by:

$$l(D|\mathbf{w}) = \sum_i^N y_i \log(\mathbf{w}^T x_i) + (1 - y_i) \log(1 - \mathbf{w}^T x_i) \quad (1)$$

where D is the data and N is the number of observations. Thus, the target unnormalised posterior log distribution is given as

$$\ln p(\mathbf{w}|D) = l(D|\mathbf{w}) + \ln p(\mathbf{w}|\alpha) + \ln q(\alpha) \quad (2)$$

where $\ln p(\mathbf{w}|\alpha)$ is the log of the prior distribution placed on the parameters given the hyperparameters, and $\ln q(\alpha)$ is the distribution of the hyperparameters. We model the parameters \mathbf{w} as having a Gaussian prior with each parameter having zero mean and its own standard deviation α_i . The α_i 's are assumed to follow a log-normal distribution with mean zero and variance 1. The α_i indicates how important the parameter associated with the input feature is. The larger the value of α_i , the more important the input feature is in predicting the audit outcomes.

The aim is to infer the parameters \mathbf{w} and hyperparameters α using Markov Chain Monte Carlo (MCMC) methods. In the literature, this problem is typically formulated as a Gibbs sampling scheme, where the hyperparameters are sampled first and then the parameters and so on [29, 31]. The approach taken in this paper is to jointly infer the parameters \mathbf{w} and hyperparameters α . This approach has the advantage of resulting in a more stable exploration of the posterior, but results in the effective parameter space being doubled—which can significantly reduce the sampling time compared to the Gibbs sampling approach.

In the following sections we present the MCMC methods used to infer the posterior distribution in Eq (2).

3.2 Metropolis-Hastings algorithm and the Metropolis Adjusted Langevin Algorithm

The Metropolis-Hastings (MH) algorithm is one of the most basic MCMC methods. The MH algorithm generates the proposed samples using a user specified proposal distribution [32, 34]. The most commonly used proposal distribution in practice is the Gaussian distribution, with the mean being the current state, and the variance being a tunable parameter [32, 34]. This in-effect creates random walk behaviour as we are simply adding Gaussian noise to the current state, with the noise being independent from the noise used to generate the current state. This random walk behaviour results in very correlated samples, particularly in high dimensions [31]. In this work, we tune the variance parameter of the MH algorithm using dual averaging [39], targeting an acceptance rate of 70% during the burn-in period.

The Metropolis adjusted Langevin algorithm (MALA) is a MCMC sampler which aims to sample from the target distribution efficiently by using first order gradient information [36, 50]. Using the first order gradient information reduces the random walk behaviour associated with the MH algorithm [36]. The MALA uses Langevin dynamics to construct the Markov chain, with the dynamics given as [36, 50]:

$$d\mathbf{w}_t = \frac{1}{2} \nabla_{\mathbf{w}} \ln \pi(\mathbf{w}) dt + dZ_t \quad (3)$$

where $\pi(\mathbf{w})$ represents the target probability density function, w represents the random variable that is to be sampled, t represents time, and Z_t is a Brownian motion process. Since this stochastic differential equation is difficult to solve analytically, the first-order Euler-Maruyama discretisation is often used to provide an approximate solution, and the solution is written as [36, 50]:

$$\mathbf{w}_{t+1} = \mathbf{w}_t + \frac{\epsilon^2}{2} \nabla_{\mathbf{w}} \ln \pi(\mathbf{w}) + \epsilon z_t \quad (4)$$

where ϵ is the step size and $z_t \sim \mathcal{N}(0, \mathbf{I})$.

The approximate solution introduces errors. In order to ensure detailed balance so that the generated chain converges to the target distribution, the Metropolis-Hasting acceptance-reject procedure is utilized. The transition probability of the MALA can be written as [50]:

$$T(\mathbf{w}'|\mathbf{w}) = \mathcal{N}(\mu(\mathbf{w}), \epsilon^2 \mathbf{I}), \quad (5)$$

$$T(\mathbf{w}|\mathbf{w}') = \mathcal{N}(\mu(\mathbf{w}'), \epsilon^2 \mathbf{I}), \quad (6)$$

$$\mu(\mathbf{w}') = \mathbf{w} + \frac{\epsilon^2}{2} \nabla_{\mathbf{w}} \ln \pi(\mathbf{w}), \quad (7)$$

$$\mu(\mathbf{w}) = \mathbf{w}' + \frac{\epsilon^2}{2} \nabla_{\mathbf{w}'} \ln \pi(\mathbf{w}'). \quad (8)$$

where $T(\mathbf{w}'|\mathbf{w})$ and $T(\mathbf{w}|\mathbf{w}')$ are transition probability distributions, \mathbf{w} is the last sample and \mathbf{w}' is the new sample generated. The final acceptance rate of the MALA takes the form:

$$\min \left[1, \frac{\pi(\mathbf{w}') T(\mathbf{w}'|\mathbf{w})}{\pi(\mathbf{w}) T(\mathbf{w}|\mathbf{w}')} \right] \quad (9)$$

A key parameter that needs to be tuned for the MALA algorithm is the step size ϵ . We tune this parameter using primal dual averaging to target an acceptance rate of 70%.

Unlike the MH algorithm, the MALA takes advantage of the gradient information of the target distribution which makes the sampler converge to the target distribution more rapidly [36, 50]. However, the generated samples are still highly correlated. In the following section we present the Hamiltonian Monte Carlo and the No-U-Turn Sampler MCMC methods which use Hamiltonian dynamics for better exploration of the target distribution.

3.3 Hamiltonian Monte Carlo and the No-U-Turn Sampler

The Hamiltonian Monte Carlo (HMC) algorithm uses first order gradient information, in a similar fashion to MALA, of the target posterior to guide its exploration of the parameter space [37, 51]. However, unlike MALA, the HMC adds an auxiliary momentum variable \mathbf{p} to the parameter space. The resultant Hamiltonian $H(\mathbf{w}, \mathbf{p})$ from this dynamic system is written

as follows [31]:

$$H(\mathbf{w}, \mathbf{p}) = U(\mathbf{w}) + K(\mathbf{p}) \tag{10}$$

where $U(\mathbf{w})$ is the negative log-likelihood of the target posterior distribution and $K(\mathbf{p})$ is the kinetic energy defined by the kernel of a Gaussian with a mass matrix \mathbf{M} [35]:

$$K(\mathbf{p}) = \frac{1}{2} \log((2\pi)^D |\mathbf{M}|) + \frac{\mathbf{p}^T \mathbf{M}^{-1} \mathbf{p}}{2}. \tag{11}$$

The trajectory of the Markov chain is driven by Hamilton's equations at a fictitious time t as follows [31]:

$$\frac{d\mathbf{w}}{dt} = \frac{\partial H(\mathbf{w}, \mathbf{p})}{\partial \mathbf{p}}; \quad \frac{d\mathbf{p}}{dt} = -\frac{\partial H(\mathbf{w}, \mathbf{p})}{\partial \mathbf{w}}. \tag{12}$$

The evolution of this Hamiltonian system must preserve both volume and total energy. Furthermore, as the Hamiltonian is separable, to traverse the space we use the leapfrog integrator [31, 37]. In the leapfrog integrator, to reach the next point in the path, we take a half step in the momentum direction, followed by a full step in the direction of the model parameters and then ending with another half step in the momentum direction [23]. The update equations for the leapfrog integration scheme are [35, 37]:

$$\begin{aligned} \mathbf{p}_{t+\frac{\epsilon}{2}} &= \mathbf{p}_t + \frac{\epsilon}{2} \frac{\partial H(\mathbf{w}_t, \mathbf{p}_t)}{\partial \mathbf{w}} \\ \mathbf{w}_{t+\epsilon} &= \mathbf{w}_t + \epsilon \mathbf{M}^{-1} \mathbf{p}_{t+\frac{\epsilon}{2}} \\ \mathbf{p}_{t+\epsilon} &= \mathbf{p}_{t+\frac{\epsilon}{2}} + \frac{\epsilon}{2} \frac{\partial H(\mathbf{w}_{t+\epsilon}, \mathbf{p}_{t+\frac{\epsilon}{2}})}{\partial \mathbf{w}}. \end{aligned} \tag{13}$$

Due to the discretisation errors arising from the leapfrog integration, a Metropolis-Hastings acceptance step is then performed in order to accept or reject the proposed sample [51], where the proposed sample parameters \mathbf{w}^* accepted with the probability [31]:

$$P(\text{accept } \mathbf{w}^*) = \min \left(1, \frac{\exp(-H(\mathbf{w}^*, \mathbf{p}^*))}{\exp(-H(\mathbf{w}, \mathbf{p}))} \right). \tag{14}$$

The overall HMC sampling process utilises a Gibbs sampling scheme, where we sample the momentum and then sample a new set of parameters given the drawn momentum. Algorithm 1 shows the pseudo-code for the HMC where ϵ is a discretisation step size. The leapfrog steps are repeated until the maximum trajectory length L is reached.

Algorithm 1: Hamiltonian Monte Carlo

Input: $N, \epsilon, L, w_{\text{init}}, H(w, p)$

Output: $(w)_{m=0}^N$

```

1:  $w_0 \leftarrow w_{\text{init}}$ 
2: for  $m \rightarrow 1$  to  $N$  do
3:    $p_{m-1} \sim \mathcal{N}(0, \mathbf{M})$ 
4:    $p_m, w_m = \text{Leapfrog}(p_{m-1}, w_{m-1}, \epsilon, L, H)$ 
5:    $\delta H = H(w_{m-1}, p_{m-1}) - H(w_m, p_m)$ 
6:    $\alpha_m = \min(1, \exp(\delta H))$ 
7:    $u_m \sim \text{Unif}(0, 1)$ 
8:    $w_m = \text{Metropolis}(\alpha_m, u_m, w_m, w_{m-1})$ 
9: end for
function Leapfrog( $p, w, \epsilon, L, H$ )

```

```

10: for  $t \leftarrow 1$  to  $L$  do
11:    $p \leftarrow p + \frac{\epsilon}{2} \frac{\partial H}{\partial w}(w, p)$ 
12:    $w \leftarrow w + \epsilon \mathcal{P}$ 
13:    $p \leftarrow p + \frac{\epsilon}{2} \frac{\partial H}{\partial w}(w, p)$ 
14: end for
   return  $-p, w$ 
function Metropolis( $\alpha_m, u_m, w_m, w_{m-1}$ )
15: if  $\alpha_m < u_m$  then
16:    $w_m = w_{m-1}$ 
17: else
18:    $w_m = w_m$ 
19: end if
   return  $w_m$ 

```

As shown in Algorithm 1, the HMC algorithm has multiple parameters that require tuning for efficient sampling, being the step size and the trajectory length. A trajectory length that is too short leads to a random walk behaviour similar to the Metropolis-Hasting method [32, 39]. A trajectory length that is too long results in a trajectory that inefficiently traces back [32, 39]. Similar conclusions can be drawn about the step size parameter. Tuning these parameters requires multiple time consuming pilot runs [32, 39].

The No-U-Turn Sampler (NUTS) automates the tuning of the leapfrog step size and trajectory length. In NUTS, the step size is tuned through primal dual averaging during the burn-in phase by targeting a specific sample acceptance rate [32, 39]. The trajectory length is tuned by iteratively doubling the trajectory length until either the chain starts to trace back or the Hamiltonian becomes infinite [32, 39, 40]. The empirical results have shown that NUTS performs at least as efficiently as and sometimes more efficiently than a well tuned standard HMC method, without requiring user intervention or costly tuning runs [39]. Thus, in this paper we use the NUTS algorithm, instead of the HMC, so that we do not perform any manual tuning of parameters.

3.4 Shadow Hamiltonian Monte Carlo

It can be shown that the leapfrog integrator only preserves the Hamiltonian up to second order [42, 44]. In order to increase accuracy, one could potentially design more accurate numerical integrators that preserve the Hamiltonian to a higher order, however, these approaches tend to be too computationally expensive [42]. Shadow Hamiltonians are perturbations of the Hamiltonian that are by design exactly conserved by the numerical integrator [38, 41], allowing one to determine the order as required.

The shadow Hamiltonian for a specific numerical integrator can be derived by performing backward error analysis on the integrator, with the shadow Hamiltonian being defined by an asymptotic expansion in the powers of the discretisation step size around the Hamiltonian [38]:

$$\tilde{H} = H + \epsilon H_2 + \epsilon^2 H_3 + \epsilon^3 H_4 + \dots \quad (15)$$

This asymptotic expansion diverges in practice, however a k^{th} order truncation of the expansion is used.

$$\begin{aligned} \tilde{H}^{[k]} &= H + \epsilon H_2 + \epsilon^2 H_3 + \epsilon^3 H_4 + \dots \\ &= \tilde{H} + \mathcal{O}(\epsilon^k) \end{aligned} \quad (16)$$

The terms H_k can be determined by matching the corresponding components of the Taylor series in terms of ϵ and the expanded exact flow of the modified differential equation of the

Hamiltonian [38]. These modified equations can be proved to be Hamiltonian for symplectic integrators such as the leapfrog [38].

In this work, we focus on a fourth-order truncation of the shadow Hamiltonian under the leapfrog integrator [38]. Since the leapfrog is second-order accurate (\mathcal{O}^2), the fourth-order truncation is conserved with higher accuracy (\mathcal{O}^4) than the true Hamiltonian [38]. The fourth-order shadow Hamiltonian for the leapfrog can be obtained by truncating the Baker–Campbell–Hausdorff (BCH) formula applied to Poisson brackets of the terms of the separable Hamiltonian [38, 41, 43, 52]:

$$\tilde{H}^{[4]} = U(\mathbf{w}) + K(\mathbf{p}) + \frac{\epsilon^2}{12} K_{\mathbf{p}}^T U_{\mathbf{w}\mathbf{w}} K_{\mathbf{p}} - \frac{\epsilon^2}{24} U_{\mathbf{w}}^T K_{\mathbf{p}\mathbf{p}} U_{\mathbf{w}} + \mathcal{O}(\epsilon^4) \tag{17}$$

where $U_{\mathbf{w}}$, $U_{\mathbf{w}\mathbf{w}}$, $K_{\mathbf{p}}$ and $K_{\mathbf{p}\mathbf{p}}$ are Jacobians and Hessians of the potential and kinetic energies, respectively. The shadow Hamiltonian in Eq (17) is non-separable in terms of \mathbf{w} and \mathbf{p} , which necessitates computational expensive momenta acceptance criteria for momenta and potential tuning of additional parameters [38, 41, 43]. This additional computational overhead is overcome by pre-processing positions and momenta before propagating through the integrator [38, 41].

The Separable Shadow Hamiltonian Hybrid Monte Carlo (S2HMC) [41] algorithm utilises a processed leapfrog integrator to create a separable Hamiltonian. The separable Hamiltonian in S2HMC is:

$$\tilde{H}(\mathbf{w}, \mathbf{p}) = U(\mathbf{w}) + K(\mathbf{p}) + \frac{\epsilon^2}{24} U_{\mathbf{w}}^T M^{-1} U_{\mathbf{w}} + \mathcal{O}(\epsilon^4) \tag{18}$$

Propagation of positions and momenta on this shadow Hamiltonian is performed after performing this reversible mapping $(\hat{\mathbf{w}}, \hat{\mathbf{p}}) = \mathcal{X}(\mathbf{w}, \mathbf{p})$ where $(\hat{\mathbf{w}}, \hat{\mathbf{p}})$ through the following fixed point iterations [38, 41]:

$$\begin{aligned} \hat{\mathbf{p}} &= \mathbf{p} - \frac{\epsilon}{24} (U_{\mathbf{w}}(\mathbf{w} + \epsilon \mathbf{M}^{-1} \hat{\mathbf{p}}) - U_{\mathbf{w}}(\mathbf{w} - \epsilon \mathbf{M}^{-1} \hat{\mathbf{p}})) \\ \hat{\mathbf{w}} &= \mathbf{w} + \frac{\epsilon^2 \mathbf{M}^{-1}}{24} (U_{\mathbf{w}}(\mathbf{w} + \epsilon \mathbf{M}^{-1} \hat{\mathbf{p}}) + U_{\mathbf{w}}(\mathbf{w} - \epsilon \mathbf{M}^{-1} \hat{\mathbf{p}})). \end{aligned} \tag{19}$$

After the leapfrog is performed this mapping is reversed using post-processing the following fixed point iterations [38]:

$$\begin{aligned} \mathbf{w} &= \hat{\mathbf{w}} - \frac{\epsilon^2 \mathbf{M}^{-1}}{24} (U_{\mathbf{w}}(\mathbf{w} + \epsilon \mathbf{M}^{-1} \hat{\mathbf{p}}) + U_{\mathbf{w}}(\mathbf{w} - \epsilon \mathbf{M}^{-1} \hat{\mathbf{p}})) \\ \mathbf{p} &= \hat{\mathbf{p}} + \frac{\epsilon}{24} (U_{\mathbf{w}}(\mathbf{w} + \epsilon \mathbf{M}^{-1} \hat{\mathbf{p}}) - U_{\mathbf{w}}(\mathbf{w} - \epsilon \mathbf{M}^{-1} \hat{\mathbf{p}})). \end{aligned} \tag{20}$$

Once the samples are obtained from S2HMC as depicted in Algorithm 2, importance weights are calculated to allow for the use of the shadow canonical density rather than the true density [38, 41]. These weights are based on the differences between the true and shadow Hamiltonians as follows:

$$b_m = \exp(-(H(\mathbf{w}, \mathbf{p}) - \hat{H}(\mathbf{w}', \mathbf{p}')) \tag{21}$$

Mean estimates of observables $f(\mathbf{w})$ which are functions of the parameters \mathbf{w} can be computed as a weighted average [38, 41].

Algorithm 2: Separable Shadow Hamiltonian Hybrid Monte Carlo

Input: $N, \epsilon, L, w_{init}, H(w, p), \hat{H}(w, p)$

```

Output:  $(w)_{m=0}^N$ , importance weights =  $(b)_{m=0}^N$ 
1:  $w_0 \leftarrow w_{\text{init}}$ 
2: for  $m \rightarrow 1$  to  $N$  do
3:    $p_{m-1} \sim \mathcal{N}(0, \mathbf{M})$ 
4:   Apply the pre-processing mapping
    $(\hat{w}, \hat{p}) = \mathcal{X}(\mathbf{w}, \mathbf{p})$ 
5:    $p_m, w_m = \mathbf{Leapfrog}(p_{m-1}, w_{m-1}, \epsilon, L, \hat{H})$ 
6:   Apply the post-processing mapping  $(\mathbf{w}, \mathbf{p}) = \mathcal{X}^{-1}(\hat{w}, \hat{p})$ 
7:    $\delta H = \hat{H}(w_{m-1}, p_{m-1}) - \hat{H}(w_m, p_m)$ 
8:    $\alpha_m = \min(1, \exp(\delta H))$ 
9:    $u_m \sim \text{Unif}(0, 1)$ 
10:   $w_m = \mathbf{Metropolis}(\alpha, u_m, w_m, w_{m-1})$ 
11:   $b_m = \exp(-(H(w_m, p_m) - \hat{H}(w_m, p_m)))$ 
12: end for

```

In this paper, we set the trajectory length for the S2HMC algorithm to 100. We then tuned the step size using dual averaging [53], targeting an acceptance rate of 70% during the burn-in phase.

4 Experimental setup

In this section we, describe the dataset used and outline the experiments undertaken.

4.1 Data description

The raw dataset was obtained from the audited financial statement data of South African municipalities over the period of 2010 to 2018 [6]. The data was sourced from the South African National Treasury website [6, 54]: <https://municipaldata.treasury.gov.za/>, with the summarised version presented in Mongwe and Malan [6]. The dataset had a total of 1 560 records, of which 55% were non-fraudulent (i.e. unqualified audit opinions). This shows that there is no large class imbalance in the data.

Table 1 provides descriptive statistics of the financial ratios, while Table 2 shows example financial ratio input features for three South African municipalities. The detailed construction of the ratios can be found in Mongwe and Malan [6], with the summary provided below:

1. *Debt to Community Wealth/Equity*—Ratio of debt to the community equity. The ratio is used to evaluate a municipality's financial leverage.
2. *Capital Expenditure to Total Expenditure*—Ratio of capital expenditure to total expenditure.
3. *Impairment of PPE, IP and IA*—Impairment of Property, Plant and Equipment (PPE) and Investment Property (IP) and Intangible Assets (IA).
4. *Repairs and Maintenance as a percentage of PPE +IP*—The ratio measures the level of repairs and maintenance relative to assets.
5. *Debt to Total Operating Revenue*—The ratio indicates the level of total borrowings in relation to total operating revenue.
6. *Current Ratio*—The ratio is used to assess the municipality's ability to pay back short-term commitments with short-term assets.
7. *Capital Cost to Total Operating Expenditure*—The ratio indicates the cost of servicing debt relative to overall expenditure.

Table 1. Five number summary of the 13 financial ratios. Note that mil represents million. Q1 and Q3 are the lower and upper quartiles. More information can be found in Mongwe and Malan [6].

Ratio	Min	Q1	Median	Q3	Max
1	-2046.45	11.97	19.43	35.49	6640.18
2	-3.81	11.22	17.85	25.36	77.66
3	-15.34	3.67	4.94	6.58	159.27
4	-1.86	0.00	0.83	1.91	170.67
5	-0.09	0.28	0.43	0.63	2.64
6	-0.88	0.56	1.16	2.19	23.06
7	-0.31	0.207	0.98	2.33	13.97
8	-147.32	-7.08	4.71	15.34	81.07
9	-67.61	26.31	32.51	41.17	159.98
10	-11.31	0.00	0.97	4.91	51.83
11	7.87	98.14	99.85	100.00	103.54
12	-114100 mil	0	0	83	1436 mil
13	-555400 mil	0	0	21	14970 mil

<https://doi.org/10.1371/journal.pone.0261245.t001>

8. *Net Operating Surplus Margin*—The ratio assesses the extent to which the entity generates operating surpluses.
9. *Remuneration to Total Operating Expenditure*—The ratio measures the extent of remuneration of the entity's staff to total operating expenditure.
10. *Contracted Services to Total Operating Expenditure*—This ratio measures how much of total expenditure is spent on contracted services.
11. *Own Source Revenue to Total Operating Revenue*—The ratio measures the extent to which the municipality's total capital expenditure is funded through internally generated funds and borrowings.
12. *Net Surplus / Deficit Water*—This ratio measures the extent to which the municipality generates surplus or deficit in rendering water service

Table 2. Example of financial ratio input features for three South African municipalities in 2010. The BUF municipality had a qualified audit opinion while CPT and ECU has unqualified audit opinions.

Ratio	BUF Municipality	CPT Municipality	ECU Municipality
1	20.72	86.73	32.95
2	17.82	20.59	14.28
3	6.73	5.66	4.89
4	2.53	9.82	2.70
5	0.50	0.53	0.58
6	2.03	1.61	1.40
7	0.00	5.13	2.45
8	11.05	9.37	4.41
9	25.95	22.74	19.48
10	0.19	8.50	3.22
11	97.77	99.71	95.68
12	48.33	85.7	-7.62
13	28.93	31.12	25.51

<https://doi.org/10.1371/journal.pone.0261245.t002>

13. *Net Surplus / Deficit Electricity*—This ratio measures the extent to which the municipality generates surplus or deficit in rendering electricity service.

4.2 Experiment description

For each of the MH, MALA, NUTS and S2HMC algorithms used in this paper, we generate five Markov chains of 10 000 samples. The first 5000 samples were used as the burn-in period, and any required tuning of algorithm parameters was performed during the burn-in period. For the NUTS and S2HMC algorithms we set the pre-conditioning matrix $\mathbf{M} = \mathbf{I}$, which is the common approach in practice [31, 33, 38].

We then assess the performance of the algorithms by generating the trace plots of the unnormalised target posterior distribution, the effective sample sizes of the generated samples, the effective sample sizes of the generated samples normalised by execution time as well as predictive performance on unseen data. Note that the execution time is the time taken to generate the samples after the burn-in period.

The ESS calculation used in this paper is the multivariate ESS metric outlined in Vats *et al.* [38, 55]. Unlike the minimum univariate ESS measure typically used to analyse MCMC results, the multivariate ESS measure of Vats *et al.* [38, 55] takes into account the correlations between the different parameter dimensions [36, 38, 55]. The minimum univariate ESS metric has the disadvantage that the estimate of the ESS ends up being dominated by the parameter dimensions that mix the slowest [38, 55]. For the S2HMC algorithm, which is an importance sampler, the multivariate ESS is adjusted by taking into account the possibility of non-uniform importance weights $(b)_{m=0}^N$ through the thinning algorithm outlined in Radivojevic *et al.* [38, 42].

The predictive performance on unseen data is performed using the accuracy measure, the receiver operating curve (ROC) as well as area under the curve (AUC). ROC plots the true positives from the model on the y-axis against the false positives on the x-axis. AUC is the area under the ROC, and represents the average miss-classifications rate. AUC is useful as a performance measure when the costs of classification are unknown, which is the case for the financial statement fraud domain [2, 21, 56, 57].

The ranking of the importance of the financial ratios is performed by calculating the mean or average α , which are the standard deviations in Eq (2), for each model parameter over the five chains. The higher the α value, the more important the input financial ratio is to the modelling of the audit outcomes.

5 Results and discussion

The experiments were implemented in PyTorch and were carried out on a 64-bit precision CPU. In evaluating the S2HMC algorithm, we set a convergence tolerance of 10^{-6} or the completion of 100 fixed point iterations.

Fig 1 shows the inference results, while Fig 2 shows the input feature importance or relevance. The results of Fig 2 are summarised, as rankings for each financial ratio, in Table 4. Table 3 shows the predictive performance of the samplers based on the Area under the Curve (AUC) and accuracy performance metrics respectively.

Fig 1(a) shows that the S2HMC produces the highest effective sampling sizes, indicating that the algorithm produces less correlated samples when compared to the other methods. NUTS has the second highest effective sample sizes, with MH and MALA having very low effective sample sizes, indicating that these two methods produce very correlated samples.

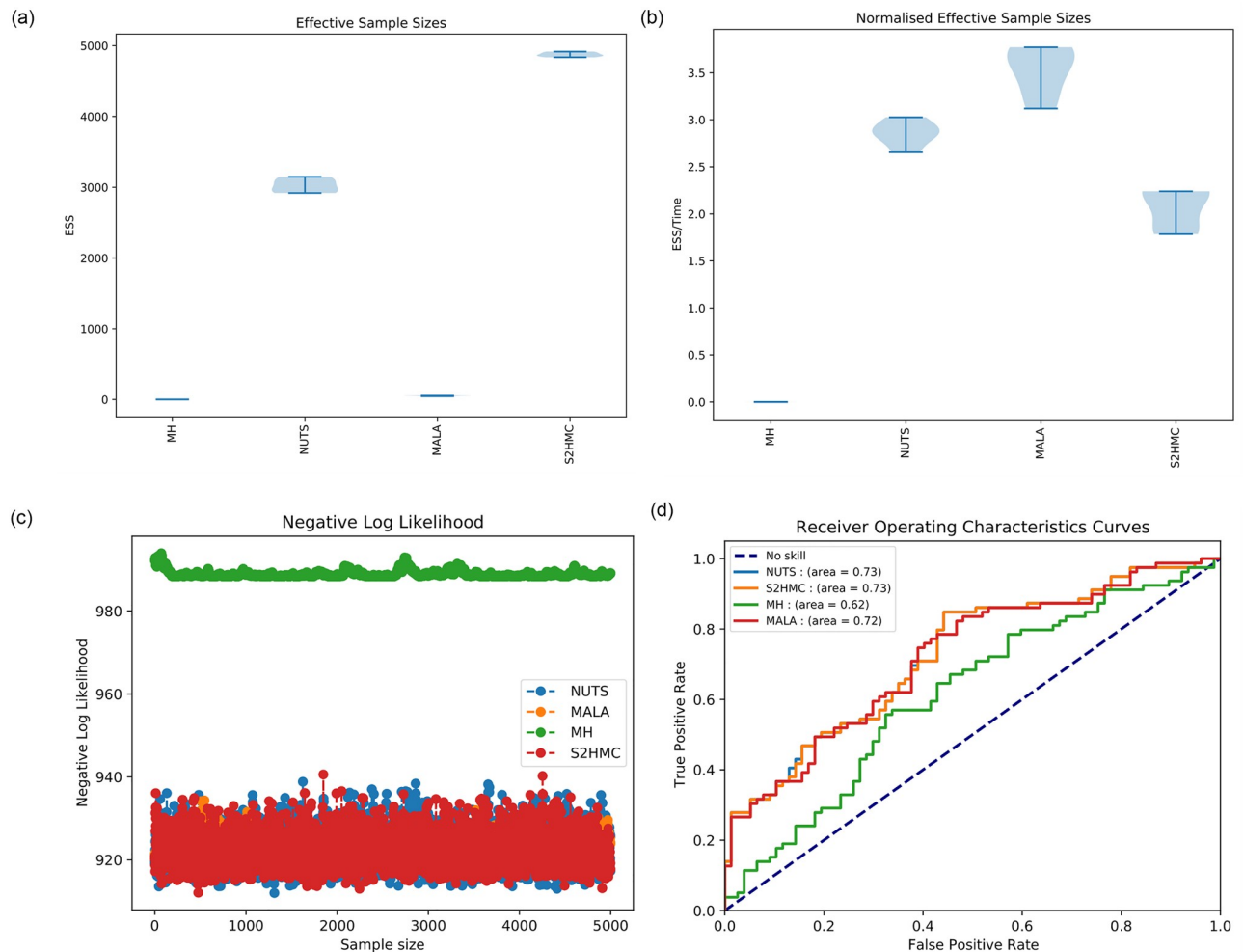


Fig 1. Inference results for the BLR-ARD model across various sampling methods. a) Effective sample sizes, b) Effective sample sizes normalised by execution time, c) Diagnostic negative log-likelihood trace plots and d) Predictive performance based on the Area under the Receiver Operating Curve.

<https://doi.org/10.1371/journal.pone.0261245.g001>

However, Fig 1(b) shows that on a normalised (by execution time) effective sample size basis, the MALA outperforms all the methods. This is testament to the fact that MALA has a very low execution time compared to NUTS and S2HMC. Although MH is also relatively fast, since the effective sample size it produces is zero, it still underperforms on a normalised effective sample size basis.

Fig 1(c) shows that the MH algorithm converges to a higher, and different negative log-likelihood than the other methods. This highlights the very poor exploration capabilities of the MH algorithm. Fig 1(d) shows that the MH algorithm has the lowest predictive performance. S2HMC and NUTS have the joint highest predictive performance, which corresponds with the high effective sample sizes generated by these methods.

Fig 2 shows the relative importance or relevance of each of the financial ratios produced by each of the MCMC methods. The results show that the MH algorithm struggles to distinguish between important and not-so-important financial ratios. This is because of the poor exploration of the target. On the other hand, the other three methods are able to extract the importance or most relevant features for the audit opinion modeling task. Table 4 shows the ranking

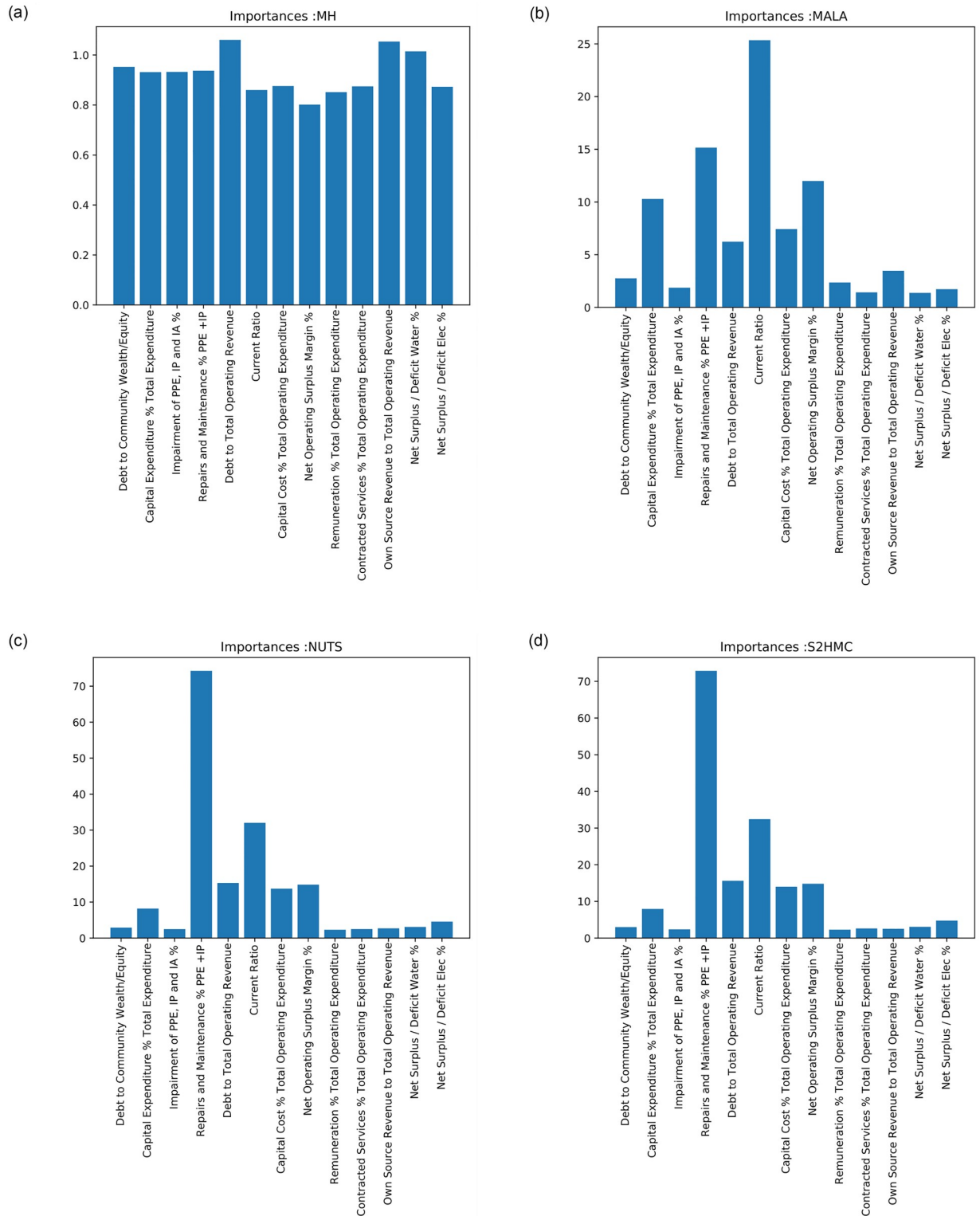


Fig 2. Mean posterior variances from each of the algorithms. The higher the value, the more important the financial ratio is to the task of modelling audit opinions. a) Importance's for MH, b) Importance's for MALA, c) Importance's for NUTS and d) Importance's for S2HMC.

<https://doi.org/10.1371/journal.pone.0261245.g002>

Table 3. Area under the receiver operating curve (AUC) and accuracy of the MCMC methods. The results were averaged over 10 runs of each algorithm.

Metric	MH	MALA	NUTS	S2HMC
AUC	0.624	0.723	0.732	0.733
Accuracy	0.651	0.737	0.744	0.753

<https://doi.org/10.1371/journal.pone.0261245.t003>

of the importance of the financial ratios produced by each of the method. The most commonly featured financial ratios in the top five ranking are:

- Ratio 4—*Repairs and Maintenance as a percentage of PPE +IP*: This ratio is selected by all four methods
- Ratio 5—*Debt to Total Operating Revenue*: This ratio is selected by the MH, NUTS and S2HMC methods
- Ratio 6—*Current Ratio*: This ratio is selected by the MALA, NUTS and S2HMC algorithms
- Ratio 8—*Net Operating Surplus Margin*: This ratio is selected by the MALA, NUTS and S2HMC algorithms
- Ratio 7—*Capital Cost to Total Operating Expenditure*: This ratio is selected by the MALA, NUTS and S2HMC algorithms.

These results are inline with those observed by Mongwe and Malan [6] for local government entities. Mongwe and Malan [6] used self-organising maps and found that the financial ratios associated with fraudulent financial statements are the current ratio, net operating surplus margin and the debt to total operating revenue. In this work, our analysis shows that the most relevant financial ratio is the repairs and maintenance as a percentage of PPE and IP, followed by the ratios found in Mongwe and Malan [6], with the fifth most relevant ratio is the capital cost to total operating expenditure ratio.

These results make intuitive sense as, for example, a high repairs and maintenance ratio means that the municipality is doing more repairs to assets than the total value of assets that it has, which is likely an indication of lack adherence to proper corporate governance as repairs to assets should typically be less than the value of those assets—else those assets should be

Table 4. Ranking of the financial ratios by each method. For example, NUTS ranks ratio 4 as the most important, while MH ranks ratio 12 as the third most important.

Ranking	MH	MALA	NUTS	S2HMC
1	5	6	4	4
2	11	4	6	6
3	12	8	5	5
4	1	2	8	8
5	4	7	7	7
6	3	5	2	2
7	2	11	13	13
8	7	1	12	12
9	10	9	1	1
10	13	3	11	10
11	6	13	10	11
12	9	10	3	3
13	8	12	9	9

<https://doi.org/10.1371/journal.pone.0261245.t004>

written-off. Furthermore, a high capital cost to total expenditure ratio means that debt repayments are the largest component of total expenditure, indicating that the entity has a large amount of debt—which might prompt the entity to act in a manner that flouts corporate governance procedures in order to hide its dire financial situation.

Mongwe and Malan [6] provide an interpretation of the current ratio, net operating surplus margin and capital cost to total operating expenditure financial ratios in terms of how they relate with audit outcomes. Our findings agree with Mongwe and Malan [6] in that we find that a high current ratio, a high net surplus operating margin and low debt to total operating revenue financial ratios are associated with entities that are less likely to engage in manipulation of their financial statements as they are in good financial standing—with the converse also being true.

These results can prove to be particularly useful for auditors as focusing on these ratios can speed up the detection of inadequate corporate governance behaviour in municipal entities, and improve the overall quality of the audits.

6 Conclusion

We present the first fully Bayesian approach to the inference of financial statement audit opinions. This Bayesian approach is applied to local government entity audit outcomes using financial ratios as inputs. The inference is performed using Metropolis-Hastings, Metropolis Adjusted Langevin Algorithm, No-U-Turn Sampler and Separable Shadow Hybrid Hamiltonian Monte Carlo algorithms. The sampling was applied to Bayesian Logistic Regression with automatic relevance determination. Automatic relevance determination (ARD) allows one to determine which features are the most important in an automated manner, and thus performing feature selection in an implicit fashion.

In this work, the parameters and the hyperparameters, which measure the relevance of the financial ratios, are jointly sampled. The results show that the Separable Shadow Hybrid Hamiltonian Monte Carlo produces the best sampling results, with the highest effective sample sizes. However, the predictive performance of the No-U-Turn Sampler and Separable Shadow Hybrid Hamiltonian Monte Carlo algorithms is found to be the same. The Metropolis-Hastings algorithm produces the worst sampling behaviour due to its random walk nature, and has both the lowest effective sample rates and predictive performance.

The results further show that the most important features in the modelling of audit outcomes for municipalities are the repairs and maintenance as a percentage of total assets ratio, current ratio, debt to total operating revenue, net operating surplus margin and capital cost to total operating expenditure ratio. This could prove to be useful for auditors as focusing on these ratios can speed up the detection of possible fraudulent behaviour of municipal entities.

This work can be improved upon by comparing the performance of the Bayesian Logistic Regression with ARD model with other models such as the Bayesian Neural Network with ARD model. Furthermore, we plan on performing this analysis for listed entities in addition to the local government entities considered in this work. The consideration of a larger set of financial ratios could also improve the results. In addition, Riemannian manifold based Markov Chain Monte Carlo methods could also be considered as they are able to better explore the target posterior distribution due to their ability to take into account the local geometry of the target distribution.

Acknowledgments

The computations in this work were performed on resources provided by the Center for High Performance Computing (CHPC) at the Council of Scientific and Industrial Research (CSIR) South Africa.

Author Contributions

Conceptualization: Wilson Tsakane Mongwe.

Data curation: Wilson Tsakane Mongwe.

Formal analysis: Wilson Tsakane Mongwe.

Funding acquisition: Tshilidzi Marwala.

Investigation: Wilson Tsakane Mongwe.

Methodology: Wilson Tsakane Mongwe, Rendani Mbuva.

Project administration: Wilson Tsakane Mongwe, Rendani Mbuva, Tshilidzi Marwala.

Resources: Wilson Tsakane Mongwe.

Software: Wilson Tsakane Mongwe.

Supervision: Tshilidzi Marwala.

Validation: Wilson Tsakane Mongwe, Rendani Mbuva.

Visualization: Wilson Tsakane Mongwe, Rendani Mbuva, Tshilidzi Marwala.

Writing – original draft: Wilson Tsakane Mongwe, Tshilidzi Marwala.

Writing – review & editing: Wilson Tsakane Mongwe, Rendani Mbuva, Tshilidzi Marwala.

References

1. Auditor-General South Africa. MFMA 2018—2019; 2020. Available from: <https://www.agsa.co.za/Reporting/MFMAReports/MFMA2018-2019.aspx>.
2. Mongwe WT, Malan KM. A Survey of Automated Financial Statement Fraud Detection with Relevance to the South African Context. *South African Computer Journal*. 2020; 32(1). <https://doi.org/10.18489/sacj.v32i1.777>
3. Cronje J. Steinhoff's market cap a mere R20bn as shares drop another 30%; 2017. Available from: <https://www.fin24.com/Companies/Retail/steinhoff-shares-drop-by-a-fifth-in-early-trade-20171220>.
4. O'Donnel J. Who's to blame for Wirecard? Germany passes the buck; 2020. Available from: <https://www.reuters.com/article/us-wirecard-accounts-responsibility/whos-to-blame-for-wirecard-germany-passes-the-buck-idUSKBN243200>.
5. Healy PM, Palepu KG. The Fall of Enron. *Journal of Economic Perspectives*. 2003; 17(2):3–26. <https://doi.org/10.1257/089533003765888403>
6. Mongwe WT, Malan KM. The Efficacy of Financial Ratios for Fraud Detection Using Self Organising Maps. In: 2020 IEEE Symposium Series on Computational Intelligence (SSCI). IEEE; 2020. p. 1100–1106.
7. Fanning KM, Cogger KO. Neural network detection of management fraud using published financial data. *International Journal of Intelligent Systems in Accounting, Finance & Management*. 1998; 7(1):21–41. [https://doi.org/https://doi.org/10.1002/\(sici\)1099-1174\(199803\)7:1%3c21::aid-isaf138%3e3.0.co;2-k](https://doi.org/https://doi.org/10.1002/(sici)1099-1174(199803)7:1%3c21::aid-isaf138%3e3.0.co;2-k)
8. Glancy FH, Yadav SB. A computational model for financial reporting fraud detection. *Decision Support Systems*. 2011; 50(3):595–601. <https://doi.org/10.1016/j.dss.2010.08.010>
9. Gupta R, Gill NS. Financial statement fraud detection using text mining. *International Journal of Advanced Computer Science and Applications*. 2012; 3(12). <https://doi.org/10.14569/ijacsa.2012.031230>

10. Kirkos E, Spathis C, Manolopoulos Y. Data Mining techniques for the detection of fraudulent financial statements. *Expert Systems with Applications*. 2007; 32(4):995–1003. <https://doi.org/10.1016/j.eswa.2006.02.016>
11. Lin JW, Hwang MI, Becker JD. A fuzzy neural network for assessing the risk of fraudulent financial reporting. *Managerial Auditing Journal*. 2003; 18(8):657–665. <https://doi.org/10.1108/02686900310495151>
12. Moepya SO, Akhoury SS, Nelwamondo FV, Twala B. The Role of Imputation in Detecting Fraudulent Financial Reporting. *International Journal of Innovative Computing, Information and Control*. 2016; 12(1):333–356. <https://doi.org/10.24507/ijicic.12.01.333>
13. Perols JL, Lougee BA. The relation between earnings management and financial statement fraud. *Advances in Accounting*. 2011; 27(1):39–53. <https://doi.org/10.1016/j.adiac.2010.10.004>
14. Persons OS. Using Financial Statement Data To Identify Factors Associated With Fraudulent Financial Reporting. *Journal of Applied Business Research (JABR)*. 2011; 11(3):38. <https://doi.org/10.19030/jabr.v11i3.5858>
15. Kiehl TR, Hoogs BK, LaComb CA, Senturk D. Evolving multi-variate time-series patterns for the discrimination of fraudulent financial filings; 2005. Available from: http://nanophobia.com/lab/cv/tom/tom-cv_files/gecco-2005-fraud.pdf.
16. Kotsiantis S, Koumanakos E, Tzelepis D, Tampakas V. Financial Application of Neural Networks: two case studies in Greece. In: *Artificial Neural Networks—ICANN 2006*. Springer Berlin Heidelberg; 2006. p. 672–681.
17. Humpherys SL, Moffitt KC, Burns MB, Burgoon JK, Felix WF. Identification of fraudulent financial statements using linguistic credibility analysis. *Decision Support Systems*. 2011; 50(3):585–594. <https://doi.org/10.1016/j.dss.2010.08.009>
18. Perols J. Financial statement fraud detection: An analysis of statistical and machine learning algorithms. *Auditing: A Journal of Practice & Theory*. 2011; 30(2):19–50. <https://doi.org/10.2308/ajpt-50009>
19. Omid P, pour Hossein N, Zeinab A. Identifying qualified audit opinions by artificial neural networks. *African Journal of Business Management*. 2012; 6(44):11077–11087. <https://doi.org/10.5897/ajbm12.855>
20. Amara I, Amar AB, Jarboui A. Detection of fraud in financial statements: French companies as a case study. *International Journal of Academic Research in Accounting, Finance and Management Sciences*. 2013; 3(3):40–51. <https://doi.org/10.6007/IJARAFMS/v3-i3/34>
21. Moepya SO, Akhoury SS, Nelwamondo FV. Applying Cost-Sensitive Classification for Financial Fraud Detection under High Class-Imbalance. In: *2014 IEEE International Conference on Data Mining Workshop*. IEEE; 2014. p. 183–192.
22. Boumediene SL. Detection and prediction of managerial fraud in the financial statements of Tunisian banks. *Accounting & Taxation*. 2014; 6(2):1–10.
23. Mbuva R, Boulkaibet I, Marwala T. Bayesian automatic relevance determination for feature selection in credit default modelling. In: *International Conference on Artificial Neural Networks*. Springer; 2019. p. 420–425.
24. Chandrashekar G, Sahin F. A survey on feature selection methods. *Computers & Electrical Engineering*. 2014; 40(1):16–28. <https://doi.org/10.1016/j.compeleceng.2013.11.024>
25. Chen YJ. On Fraud Detection Method for Narrative Annual Reports. In: *The Fourth International Conference on Informatics & Applications (ICIA2015)*; 2015. p. 121–129. Available from: https://www.researchgate.net/publication/280222998_Proceedings_of_The_Fourth_International_Conference_on_Informatics_Applications_Takamatsu_Japan_2015.
26. Song XP, Hu ZH, Du JG, Sheng ZH. Application of Machine Learning Methods to Risk Assessment of Financial Statement Fraud: Evidence from China. *Journal of Forecasting*. 2014; 33(8):611–626. <https://doi.org/10.1002/for.2294>
27. Chen S. Detection of fraudulent financial statements using the hybrid data mining approach. *Springer-Plus*. 2016; 5(1):1–16. <https://doi.org/10.1186/s40064-016-1707-6> PMID: 26848429
28. Fernández-Gómez MA, García-Lagos F, Sánchez-Serrano JR. Integrating corporate governance and financial variables for the identification of qualified audit opinions with neural networks. *Neural Computing and Applications*. 2015; 27(5):1427–1444. <https://doi.org/10.1007/s00521-015-1944-6>
29. Mbuva R, Boulkaibet I, Marwala T. Automatic Relevance Determination Bayesian Neural Networks for Credit Card Default Modelling; 2019.
30. Salimans T, Kingma D, Welling M. Markov chain monte carlo and variational inference: Bridging the gap. In: *International Conference on Machine Learning*. PMLR; 2015. p. 1218–1226.
31. Neal RM. Bayesian learning via stochastic dynamics. In: *Advances in neural information processing systems*; 1993. p. 475–482.

32. Mbuvha R, Marwala T. Bayesian inference of COVID-19 spreading rates in South Africa. *PloS one*. 2020; 15(8):e0237126. <https://doi.org/10.1371/journal.pone.0237126> PMID: 32756608
33. Betancourt M. A conceptual introduction to Hamiltonian Monte Carlo. arXiv preprint arXiv:170102434. 2017;.
34. Hastings WK. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*. 1970; 57(1):97–109. <https://doi.org/10.1093/biomet/57.1.97>
35. Neal RM. Bayesian learning for neural networks. vol. 118. Springer Science & Business Media; 2012.
36. Girolami M, Calderhead B. Riemann manifold langevin and hamiltonian monte carlo methods. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*. 2011; 73(2):123–214. <https://doi.org/10.1111/j.1467-9868.2010.00765.x>
37. Duane S, Kennedy AD, Pendleton BJ, Roweth D. Hybrid monte carlo. *Physics letters B*. 1987; 195(2):216–222. [https://doi.org/10.1016/0370-2693\(87\)91197-X](https://doi.org/10.1016/0370-2693(87)91197-X)
38. Mongwe WT, Mbuvha R, Marwala T. Antithetic Magnetic and Shadow Hamiltonian Monte Carlo. *IEEE Access*. 2021; 9:49857–49867. <https://doi.org/10.1109/ACCESS.2021.3069196>
39. Hoffman MD, Gelman A. The No-U-turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research*. 2014; 15(1):1593–1623.
40. Afshar HM, Oliveira R, Cripps S. Non-Volume Preserving Hamiltonian Monte Carlo and No-U-Turn-Samplers. In: *International Conference on Artificial Intelligence and Statistics*. PMLR; 2021. p. 1675–1683.
41. Sweet CR, Hampton SS, Skeel RD, Izaguirre JA. A separable shadow Hamiltonian hybrid Monte Carlo method. *The Journal of chemical physics*. 2009; 131(17):174106. <https://doi.org/10.1063/1.3253687> PMID: 19894997
42. Radivojević T, Akhmatskaya E. Mix & Match Hamiltonian Monte Carlo. arXiv preprint arXiv:170604032. 2017;.
43. Izaguirre JA, Hampton SS. Shadow Hybrid Monte Carlo: An Efficient Propagator in Phase Space of Macromolecules. *J Comput Phys*. 2004; 200(2):581–604. <https://doi.org/10.1016/j.jcp.2004.04.016>
44. Akhmatskaya E, Reich S. The targeted shadowing hybrid Monte Carlo (TSHMC) method. In: *New Algorithms for Macromolecular Simulation*. Springer; 2006. p. 141–153.
45. Schilit HM. Financial Shenanigans: Detecting Accounting Gimmicks That Destroy Investments (corrected November 2010). *CFA Institute Conference Proceedings Quarterly*. 2010;27(4):67–74. <https://doi.org/10.2469/cp.v27.n4.1>
46. Donnelly L. PIC still limping from Steinhoff blow; 2018. Available from: <https://mg.co.za/article/2018-06-08-00-pic-still-limping-from-steinhoff-blow>.
47. Bratton WW. Enron and the Dark Side of Shareholder Value. *SSRN Electronic Journal*. 2002;. <https://doi.org/10.2139/ssrn.301475>
48. Auditor-General South Africa. Background to the three aspects we audit; 2011. Available from: https://www.agsa.co.za/portals/0/AGSA_Terminology.pdf.
49. Şen IK, Terzi S. Detecting falsified financial statements using data mining: empirical research on finance sector in Turkey. *Maliye Finans Yazilari*. 2012; 26(96):67–82.
50. Gu M, Sun S. Neural Langevin Dynamical Sampling. *IEEE Access*. 2020; 8:31595–31605. <https://doi.org/10.1109/ACCESS.2020.2972611>
51. Neal RM, et al. MCMC using Hamiltonian dynamics. *Handbook of Markov Chain Monte Carlo*. 2011; 2(11):2.
52. Barp A, Briol FX, Kennedy AD, Girolami M. Geometry and dynamics for Markov chain Monte Carlo. *Annual Review of Statistics and Its Application*. 2018; 5:451–471. <https://doi.org/10.1146/annurev-statistics-031017-100141>
53. Hoffman MD, Gelman A. The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo; 2011.
54. National Treasury Republic Of South Africa. Municipal Finance Data; 2019. Available from: <https://municipaldata.treasury.gov.za/>.
55. Vats D, Flegal JM, Jones GL. Multivariate output analysis for Markov chain Monte Carlo. *Biometrika*. 2019; 106(2):321–337. <https://doi.org/10.1093/biomet/asz002>
56. Bradley AP. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern recognition*. 1997; 30(7):1145–1159. [https://doi.org/10.1016/S0031-3203\(96\)00142-2](https://doi.org/10.1016/S0031-3203(96)00142-2)
57. Gaganis C. Classification techniques for the identification of falsified financial statements: a comparative analysis. *Intelligent Systems in Accounting, Finance & Management*. 2009; 16(3):207–229. <https://doi.org/10.1002/isaf.303> PMID: 33152585