Detection of Data Manipulation Using Deep Learning

Poj Netsiri
Immatriculation number: 12402153
e124021538@student.tuwien.ac.at

Main Supervisors
Assoz.Prof. PD Dr. Jakob Müllner, Wirtschaftsuniversität Wien
Asst.Prof. Dr. Harald Puhr, Universität Innsbruck
Co-Supervisor
Assoz.Prof. Dr. Nysret Musliu, Technische Universität Wien

1 Abstract

Unethical data manipulation involves the deliberate falsification or selective alteration of data to support a specific hypothesis, leading to fraudulent scientific conclusions. Such misconduct can result in serious consequences, including research paper retractions, loss of credibility, withdrawal of funding, and potential harm to society.

This project aims to develop Deep Learning (DL) modes capable of detecting and classifying manipulated datasets. By training algorithms on a combination of authentic and synthetically manipulated data, the system will learn to identify anomalies and distinguish between genuine and fraudulent submissions. The project contributes to improving transparency and upholding integrity in scientific research.

2 Introduction

Unethical data practices threaten the credibility of research by presenting manipulated outcomes as legitimate findings. This manipulation can be motivated by the desire for academic recognition, funding, or institutional pressure. Common forms include:

- **Data falsification**: Altering or deleting data to better fit hypotheses.
- **P-hacking**: Repeatedly testing data until significant p-values are obtained.

Such practices distort the scientific record, misguide policy decisions, and erode public trust. Detecting these patterns using data-driven methods can play a vital role in research integrity.

3 Methodology and Workflow

The methodology integrates statistical fraud detection with a Deep Learning pipeline:

1. **Data Collection**: Authentic financial data (Vorarlberg municipal finance statistics[1]) is collected as the base dataset for experimentation.
2. **Data manipulation**: Three hypotheses are formed. Common techniques such as falsification are utilized to manipulate the data to fit one of these hypotheses.
3. **Data Exchange and Collaboration**: Exchange manipulated, raw datasets and hypothesis with a collaborator (Tuvshin Selenge).
4. **Conventional Approach**: Apply traditional methods, including histogram analysis and Benford's Law, to detect digit-level irregularities and other anomalies.
5. **DL Approach**: Design and train an unsupervised neural network (e.g., anomaly detection model) to learn latent patterns distinguishing authentic from manipulated data. The model is optimized for sensitivity to non-obvious irregularities.
6. **Comparison and Evaluation**: Evaluate using metrics such as precision, recall, and F1-score. Analyze trade-offs between conventional techniques and deep learning approaches in detecting complex or subtle data tampering.

This project distinguishes itself by combining a theoretically grounded method (Benford's Law) with the adaptability of modern deep learning, forming a dual strategy for validating hypotheses and screening data fraud.

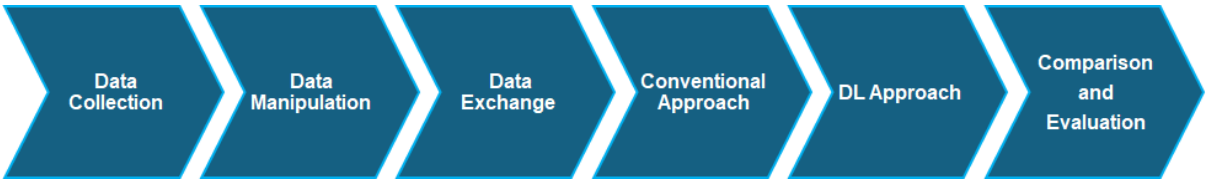

Figure 1: Workflow

4 Expected Results

The project is expected to produce a robust, scalable DL-based detection model capable of identifying manipulated datasets with high accuracy. The tool can support researchers, reviewers, and funding agencies in early-stage auditing by flagging anomalous data submissions. Comparative analysis will also highlight the strengths and limitations of deep learning versus traditional statistical methods, providing a roadmap for further development in automated scientific fraud detection.

5 References

1. Vorarlberg Municipal finance statistics up to 2019. `https://www.data.gv.at/katalog/dataset/7e322fe6-e881-42cb-9f85-12364d22b2b3#resources`. Accessed: 2025-04-12.
2. Data Falsificad (Part 1):"Clusterfake". `https://datacolada.org/109`. Accessed: 2025-04-01.
3. Reinhart and Rogoff are wrong about austerity. `https://peri.umass.edu/wp-content/uploads/joomla/images/wp322FT_Pollin_Ash.pdf`. Accessed: 2025-04-01.