

**Erickson Mbuvi**  
**Data Scientist | Software Developer**

+254700120294 | [mutisyaerickson@gmail.com](mailto:mutisyaerickson@gmail.com) | [www.linkedin.com/in/erickson-mbuvi-data-scientist](https://www.linkedin.com/in/erickson-mbuvi-data-scientist) |  
<https://github.com/ericks-on> | Kenya

## SUMMARY

---

Data Scientist with expertise in machine learning, predictive analytics, and data management. Skilled in developing and deploying data-driven solutions, improving business decision-making, and managing data quality and governance. Adept at leveraging advanced statistical techniques and big data platforms to solve complex business challenges.

## SKILLS

---

- **Programming & Frameworks:** Python, TensorFlow, Keras, Scikit-learn, SQL, Flask, NextJS, TypeScript, Tailwind CSS
- **Data & Visualization:** Power BI, Tableau, Seaborn, Matplotlib, Data Analysis, Data Preprocessing, A/B Testing, deep learning, dashboard development, data quality testing
- **Web Development:** HTML, CSS, Vanilla JavaScript, jQuery
- **Databases:** MySQL, PostgreSQL, Big Data Tools (e.g., Hadoop, Spark)
- **Data Management:** Data Quality Testing, Metadata Management, Master Data Management, Data Privacy & Security, Data Governance
- **Soft Skills:** Problem-Solving, Communication, Team Collaboration, Adaptability, Analytical Thinking, Time Management, Continuous Learning

## Projects

---

### Classification of Online Shoppers - [https://github.com/ericks-on/predicting\\_purchase\\_intent](https://github.com/ericks-on/predicting_purchase_intent)

Developed a high-performing classification model to predict customer purchase intent.

- Built models like logistic regression and decision trees for accurate predictions.
- Conducted feature engineering, handled missing values, and optimized model performance through hyperparameter tuning.
- Delivered insights on customer behavior for strategic recommendations and prepared non-technical presentation slides.

### NLP: AI-Generated text detection - [https://github.com/ericks-on/NLP-Ai\\_generated\\_Text\\_detection](https://github.com/ericks-on/NLP-Ai_generated_Text_detection)

Created a machine learning model to detect AI-generated text using NLP.

- Tested multiple models, including Logistic Regression, Random Forest, and Neural Networks, selecting the optimal approach.
- Integrated the model into a Flask-based web application for practical deployment.
- Leveraged tools such as Scikit-learn, NLTK, Pandas, and TensorFlow for end-to-end model development.

### Fraud Detection System for Insurance Claims - <https://github.com/ericks-on/insurance-fraud-detection>

- Developed a Logistic Regression model fine-tuned through Grid Search, achieving 80% accuracy and 70% recall.
- Conducted initial deep learning experiments using TensorFlow and Keras to benchmark performance.
- Collaborated on model development and optimization with team members, focusing on identifying fraudulent claims.
- Built an Admin Dashboard to monitor claims, detect suspicious activities, and blacklist fraudulent users.
- Designed a User Dashboard to simulate claim filing while enforcing restrictions for blacklisted users.
- Delivered actionable insights through Tableau dashboards, visualizing trends and uncovering patterns in the dataset.
- Worked with cross-functional team members, including data engineers and analysts, to ensure an end-to-end solution.

EXPERIENCE

---

Python Developer | Ervin Solutions

January 2022 – Present.

- Developed Python scripts to automate repetitive tasks, such as data cleaning, report generation, and file management, significantly improving efficiency and reducing manual effort.
- Designed and implemented RESTful APIs using Flask to enable seamless data exchange between applications and services.
- Built web applications using Flask, integrating front-end technologies like HTML, CSS, and JavaScript to create dynamic and user-friendly interfaces.
- Utilized Python libraries like Pandas and NumPy to analyze large datasets, extract meaningful insights, and visualize trends.
- Developed and deployed advanced econometric and machine learning models for predictive analytics, classification, and clustering.
- Designed and implemented end-to-end data pipelines to ensure high-quality data management, including data preprocessing, cleansing, and governance.
- Enhanced business outcomes by generating actionable insights through predictive modeling, customer profiling, and segmentation analysis.
- Built dashboards in Tableau and Power BI for visualizing trends, key performance indicators, and ad-hoc analytics.
- Collaborated with cross-functional teams to integrate data privacy and security measures into the analytics workflows.
- Leveraged Python and big data platforms like Hadoop and Spark to manage large datasets for time-series analysis and forecasting.

Junior Developer | Enta Studios

February 2023 – August 2023 | Nairobi, Kenya

- Collaborated with UI/UX teams to build an engaging and intuitive user interface for a cutting-edge movie streaming platform.
- Contributed to feature development and improved the user experience by implementing feedback-driven enhancements and testing.
- Designed and implemented a centralized data management system for tracking user behavior and content performance on a movie streaming platform.
- Collaborated with the data engineering team to improve data collection processes, ensuring clean and actionable datasets for analysis.
- Conducted exploratory data analysis to identify patterns in customer engagement and provide data-driven recommendations for enhancing user experience.
- Supported the creation of business intelligence dashboards, enabling stakeholders to monitor platform metrics in real-time.
- Contributed to data retention and archiving processes to ensure compliance with industry standards.
- Streamlined data workflows, improving the efficiency of analytics projects by 30%.

EDUCATION

---

<b>Multimedia University of Kenya</b>   Bachelor of Science in Mathematics and Computer Science	Sept 2020 - Nov. 2024
<ul style="list-style-type: none"><li>• Second class upper division</li></ul>	
<b>Moringa School</b>   Data Science	May 2024 - Nov. 2024
<ul style="list-style-type: none"><li>• Certificate</li></ul>	

## References

---

**Antonny Muiko** | Technical mentor, Moringa School

- [Antonny.muiko@moringaschool.com](mailto:Antonny.muiko@moringaschool.com)

**Ian Ndirangu** | CEO, Enta Studios

- [linkedin.com/in/ian-ndirangu-0b6492223](https://linkedin.com/in/ian-ndirangu-0b6492223)