




# Kingston Teffy Roy Deivapathinathan

 [kingstondroy@gmail.com](mailto:kingstondroy@gmail.com)  [github.com/kingstonDRoy](https://github.com/kingstonDRoy)  [www.linkedin.com/in/kingstonteffyroy](https://www.linkedin.com/in/kingstonteffyroy)

 +49 1793960948  Kossaettenweg 25, 14476 Potsdam

**Work Permit:** Authorised to Work in Germany

## PROFESSIONAL SUMMARY

Aspiring Data Analyst with a passion for uncovering insights through data visualization, statistical analysis, and BI. Proficient in Python, SQL, Power BI, and Tableau. Experienced in turning complex data into actionable insights to drive decisions. Currently exploring A/B testing, data storytelling, and dashboard development to optimize processes and identify trends for business impact.

## EDUCATION

**GISMA University of Applied Sciences, Potsdam, Germany**

**09/2022 - 12/2023**

MSc in Data science, AI and Digital Business | Grade: 2.4

- Elected Student Representative | 2022 – 2023
- Master's Thesis – Ethics of AI-Generated Content in Social Media

**ANNA University, Chennai, India**

**08/2014 - 04/2018**

B.E., Computer Science and Engineering | 7.38 CGPA

- Computer Science Department Co-ordinator | 2016 - 2018

## PROFESSIONAL EXPERIENCE

**Operations Associate, Amazon Logistics GmbH, Werder, Germany**

**10/2023 - 04/2025**

- Promoted within 3 months to the Problem-Solving Team, where I ensured data accuracy and streamlined reporting by aligning cross-department efforts using Asana, Chime, and Amazon WMS tools.
- Strengthened cross-functional collaboration and boosted efficiency by advancing German language proficiency, enabling seamless communication in a dynamic bilingual workplace.

**Deep Learning Internship, iNeuron.ai, India**

**05/2022 - 09/2022**

- Leveraged NLP techniques and advanced machine learning algorithms (spaCy, NLTK) to preprocess, extract features, and analyze large-scale textual data for an automated plagiarism detection system, increasing accuracy by 15% and improving operational efficiency by 20%.
- Applied TF-IDF, Cosine Similarity, cross-validation, and hyperparameter tuning to optimize text comparison models; collaborated in Agile teams to build scalable data pipelines and contribute to PoCs, enhancing data processing workflows and stakeholder reporting.

**Data Science Tutor, Freelance, India**

**07/2020 - 05/2022**

- Mentored non-technical learners in Python for Data Science, with a focus on applied analytics, machine learning foundations, and hands-on project development using tools like Jupyter, Google Colab, and VS Code.
- Delivered instruction on data wrangling, ETL processes, web scraping, and exploratory data analysis (EDA); trained students in data visualization using Tableau, Power BI, Matplotlib, and Seaborn to build interactive dashboards and generate actionable insights.

**Data Science Internship, Rubix - AI Company, Bangalore, India**

**12/2019 - 06/2020**

- Conducted employee performance analysis and inventory forecasting using regression, time-series, and ensemble models; applied TensorFlow, XGBoost, and PySpark for large-scale data preprocessing, feature engineering, and model optimization, improving forecasting accuracy and operational efficiency.
- Delivered predictive analytics and business intelligence insights to support strategic decision-making; earned the Certified Data Scientist credential from IABAC Netherlands, with specialization in data-driven performance optimization and KPI reporting.

- Contributed to application development using Odoo ERP, integrating payment gateways and applying core python programming concepts to enhance system functionality and streamline transactional workflows.
- Designed and developed a Blood Bank Management System to automate donor and inventory management, resulting in improved process efficiency and real-time tracking; earned top-grade (A) for project execution and presentation.

## LANGUAGES

- **English** - Native
- **German** – B1 (Currently B2)

## PROJECTS [🔗 \(Click here\)](#)

- **Employee Performance Prediction – Workforce Analytics:** Engineered predictive models using XGBoost, LightGBM, Random Forest, with SMOTE, PCA, SHAP, Bayesian Optimization for classification, feature selection, and KPI forecasting. Automated versioning and reproducibility with MLflow, DVC, and Hyperopt; built interactive reports in Power BI and Jupyter Notebooks for HR performance analysis, workforce planning, and data governance. **Tools:** Python, Scikit-learn, Power BI, MLflow, DVC, Hyperopt, SHAP, SMOTE, Jupyter.
- **Customer Churn Prediction – Retention Strategy, Predictive BI, Dashboards:** Designed end-to-end churn prediction pipeline using Optuna, Stratified K-Fold, Target Encoding, and SHAP for model interpretability and KPI insights. Created dynamic Streamlit dashboards for marketing analytics, including churn scores, customer segmentation, and campaign impact tracking, enhancing data-driven retention strategies and automated insight delivery. **Tools:** Python, XGBoost, LightGBM, Optuna, SHAP, Streamlit, Scikit-learn, Jupyter.
- **NLP Pipeline Deployment – CI/CD, Cloud Analytics, Model Serving:** Deployed real-time text classification and sentiment analysis pipelines using Hugging Face Transformers, FastAPI, ONNX Runtime, containerized with Docker and orchestrated via Kubernetes. Set up CI/CD workflows using GitHub Actions, Prometheus, TensorBoard, and Weights & Biases for model monitoring, performance logging, and production deployment on AWS and Google Cloud Platform (GCP). **Tools:** Transformers, FastAPI, Docker, Kubernetes, ONNX, AWS, GCP, Prometheus, TensorBoard, Weights & Biases, GitHub Actions.

## TECHNICAL STACK & PROFESSIONAL SKILLS

- **Data Analysis & Visualization:** Tableau, Power BI, Looker Studio, Matplotlib, Seaborn, Plotly, Datawrapper, Excel (Advanced), PivotTables, Google Sheets (AppScript)
- **Data Management & Processing:** SQL, dbt, Apache Spark, PySpark, BigQuery, ETL/ELT Pipelines, Airflow, Data Cleansing, Data Wrangling, Web Scraping (BeautifulSoup, Selenium)
- **Programming & Scripting:** Python (for analytics, automation, and reporting), Jupyter Notebooks, Google Colab, VS Code, Git/GitHub, Bash (basic)
- **Statistical & Analytical:** A/B Testing, Hypothesis Testing, Time Series Analysis, Regression Modeling, Cluster Analysis, Feature Engineering, Outlier Detection, Correlation Analysis
- **Business Intelligence & Reporting:** Dashboard Design, KPI Reporting, Business Metrics, Ad-Hoc Reporting, Automated Reporting Pipelines, Stakeholder Communication
- **Cloud & Data Platforms:** Google Cloud Platform (BigQuery, Cloud Functions), Microsoft Azure, Snowflake, AWS (Athena, S3 for analytics use cases)
- **Tools & Workflow:** Notion, Confluence, Jira, Trello, Slack, Microsoft Office Suite, Agile/Scrum
- **Certifications & Methodologies:** Certified Data Scientist – IABAC Netherlands, Business Intelligence, Predictive Analytics, Data-Driven Decision Making
- **Soft Skills:** Analytical Thinking, Problem Solving, Communication with Non-Technical Stakeholders, Cross-Functional Collaboration, Data Storytelling

## CERTIFICATIONS

- Complete MLOps Bootcamp - Udemy
- Data Intelligence Job Simulation - PwC, Switzerland
- Certified Data Scientist - IABAC, Netherlands