

# Rasika Edirisinghe

Status [Marie Curie Fellow](#) - University Cork College, Ireland  
 PHD in Computer Science (Reading) - University Cork College, Ireland

Contact [rasikahendrix@gmail.com](mailto:rasikahendrix@gmail.com) | +353872666627

LinkedIn <https://www.linkedin.com/in/rasika-chamara/>

GitHub <https://mrhendrixsl.github.io/>  
<https://github.com/mrHendrixSL/>



## Summary & Research Interests

NLP | Machine Learning | Deep Learning | LLMs | Computational Linguistics

I'm an AI researcher specializing in Natural Language Processing, Machine Learning, and Transformers, with a strong foundation in applied data visualization. Currently pursuing a fully funded PhD in Computer Science at University College Cork, Ireland, under the Horizon Europe MSCA program, my research focuses on developing LLM-powered text reuse detection systems. I bring over four years of cross-functional industry experience as a data analyst and implementation engineer, building cloud-based BI tools and automation pipelines. I thrive at the intersection of research and real-world application, aiming to deliver explainable, human-centered AI solutions for language and data-intensive environments.

## Education

Sep 2024 - Sep 2027	<b>PHD - Computer Science (Reading) - Artificial Intelligence &amp; Computational Linguistics</b> ( <i>Thesis: Language Modelling and Visualization for Parallel Text Analysis.</i> )  <i>Designing advanced text reuse detection systems for diachronic corpora and low-resource textual domains using Large Language Models (LLMs) and interactive visual analytics.</i> Addressing challenges such as orthographic variation, semantic drift, and context-dependent alignment across time and textual use cases.	University Cork College, Cork, Ireland
Sep 2021 - Sep 2023	<b>Master of Science - MS, Business Analytics (Final Grade - Distinction)</b>  <b>Thesis - Twitter Sentiment Reason Mining Framework to Identify Major Problems in US Healthcare Industry.</b> (Python based seamless NLP framework, Final Grade - A)	Robert Gordon University, Aberdeen, UK
Nov 2016 - Oct 2018	<b>Bachelor of Science with Honors - BS Hons, Business information Systems</b> <b>Thesis - Computer Aided Diagnosis System for Lung Cancer Detection (Standalone Java Based CADE system)</b>  <i>Modules - Analytics &amp; Business Intelligence, Cyber Security, Professional &amp; Ethical Issues in IT, Information Systems Project Management</i>	Cardiff Metropolitan University, Cardiff, UK
May 2014 - May 2016	<b>BTEC Higher National Diploma - BTEC HND, Software Development</b> <b>BTEC Higher National Diploma - BTEC HND, Computing and Systems Development (QCF)</b>  <i>Modules - Information Systems, Basic of Computing, Basic of Programming, System Analysis &amp; Design, Data Structures and Algorithms, Object Oriented Programming, Procedural Programming, Data Analysis and Design, Java Application Development, Programming In .Net, Data Structures with Java, Database Development, Web Applications Development, Networking Technologies, Business Skills for E-Commerce, IT SECURITY FOR Organizations, Employability and Professional Development</i>	Pearson Edexcel, UK

## Academic Experience

Aug 2023 -  
Current

Visiting Lecturer - Department of Electrical & Information Engineering  
Conducting lectures on

University of  
Ruhuna

1. Introduction to Data Visualization with PowerBI

*Data Sources, Building Visualizations in PowerBI, Visualization principles, Color theory, Gestalt principles of visual perception, Advanced Data Visualization Techniques, Calculated fields and measures, hierarchies and drill-down visualizations, DAX expressions, Dashboard Creation and Sharing, building interactive dashboards, designing effective layouts and navigation, publishing and sharing.*

2. Excel for Data Science

*Data Sources, Formats & Cleaning, Data Visualization & Identifying Patterns, Data Transformation, Formulas, Functions & Conditional Statements, PivotTables Basics & Data Models, PivotCharts Introduction, Time Series Analysis with Excel, Statistical Analysis & hypotheses Testing with Ms Solver, implementing basic machine learning algorithms, building interactive data dashboards, Spreadsheet based Business Modelling & simulations, Linear programming with Solver.*

Aug 2023 -  
Current

Visiting Lecturer - Sri Lanka Telecom Training Centre  
Conducting lectures on

Sri Lanka  
Telecom PLC

1. Fundamentals of Artificial Intelligence (AI) & Intelligent Systems - Unit Code - K/618/5660 (Pearson BTEC)

*AI fundamentals, AI types, AI applications, Approaches to AI, AI tools and techniques, Image recognition, NLP, conversational AI, AI analysis, Ethics of AI, Challenges of AI, Opportunities of AI.*

2. Big Data & Visualization - Unit Code - J/618/6251 (Pearson BTEC)

*Big data and visualization for decision making, Statistical and graphical techniques, tools and industry software solutions for big data and visualization, Data-driven decision making, Python, jupyter notebooks, Ms Power BI, DAX, Roles, responsibilities, and challenges for data specialists, compliance to policies, laws on data privacy, security, data source challenges & data governance.*

3. Business Intelligence - Unit Code - M/615/1641 (Pearson BTEC)

*Business processes & capturing activities, information flows, and decision points within an organization, Business analysis, Business intelligence systems, databases, data warehouses, OLAP, data visualization. Tools for business intelligence, descriptive analytics, predictive analytics Solutions for business intelligence, automate procedures, Legal, regulatory, ethical constraints, data privacy, security.*

Configure and troubleshoot software applications and practical environments on lab computers to ensure optimal conditions for training sessions.

Preparation course works /tests & grading of students in accordance with Pearson's standards and guidance.

## Industry Experience

May 2023 -  
Current

Data Analyst

EFL Global

- *Model, document, develop, test, deploy and maintain data wrangling and analytical automation scripts / UI level Apps. (Python, VBA, DAX).*
- *Modelling | Implementing | Maintaining Azure Databricks backend for Automations | Reporting | PowerBI | Ms Excel.*
- *Business workflow automation | Ms Power Automate | Ms Logic Apps.*
- *Ms Power BI | Tableau.*
- *Azure Analysis Services | Azure Data Lakes | Azure Blob Storage. | Azure Synapse Analytics.*

Nov 2022 – May 2023	<b>Implementation Engineer</b> <ul style="list-style-type: none"> <li>Implementing and maintaining supply chain BI solutions.</li> <li>Collaborated with stakeholders to gather requirements and conceptualize initial technical solutions, translating business needs into actionable prototypes and system designs as a Technical BA</li> <li>Providing training and support to clients in reporting and analytics.</li> </ul>	Affno Asia Pacific
Sep 2019– Nov 2022	<b>Operations Analyst</b> <ul style="list-style-type: none"> <li>Doctrix. (Qlik-Sense cloud based Self-Service RCM BI platform) (Jan 2020 Up to Nov 2022)</li> <li>KPI   Industry Benchmarks   SLA Monitoring.</li> <li>Data Steward. (In charge of Data Sanity)</li> <li>Process trainer. (Multiple Clients, 2021 Mar Up to Nov 2022)</li> <li>Reporting Automation. (Python +VBA). (2021 Mar up to Nov 2022)</li> <li>QlikView   Amazon DynamoDB.   Amazon Kinesis Data Streams.</li> </ul>	SYNERGEN Health

## Internships

Mar 2019 – Sep 2019	<b>Associate Operations Analyst</b> <ul style="list-style-type: none"> <li>Supports data management processes.</li> <li>Support trade processes.</li> <li>client reporting.</li> <li>Process Monitoring.</li> <li>Data Sanity checks.</li> <li>Business process monitoring.</li> </ul>	SYNERGEN Health
Sep 2018 – Mar 2019	<b>Intern – Business Analyst</b> <ul style="list-style-type: none"> <li>Assist requirement gathering using interviews, document analysis, requirements workshops, surveys, site visits.</li> <li>Assist Business process descriptions, use cases, scenarios, task, and workflow analysis.</li> <li>Assist translation of reporting requirements into formats accepted by Trax service providers and oversee the timely development process.</li> <li>Assist in Development of requirement specifications according to standard templates, using natural language.</li> <li>Support pre-sales activities.</li> </ul>	Global System Solutions

## Presentations & Academic Contributions

Jun 2023	Presented a research paper titled “Sentiment Reason Mining Framework for Analyzing Twitter Discourse on Critical Issues in US Healthcare Industry” at the INTERNATIONAL CONFERENCE ON SMART COMPUTING AND SYSTEMS ENGINEERING 2023.	
Oct 2023	Contributed to the System Development Project Poster Presentation held by the Department of Industrial Management, Faculty of Science, University of Kelaniya as an evaluator.	
Aug 2023	<p>Contributed to the curriculum development of the Foundation to Data Science Unit module, - Department of Electrical &amp; Information Engineering - University of Ruhuna as an external consultant/reviewer regarding,</p> <ol style="list-style-type: none"> <li>Introduction to python with google colab/jupyter notebooks.</li> <li>Introduction to preprocessing, data acquisition and reliability, import libraries, import data.</li> <li>Cleaning data, handling missing data, outlier detection, one hot encoding, handling categorical variables.</li> </ol>	

4. Dimensionality reduction (PCA, ICA, etc).
5. Data splitting and cross validation, training/validation/test sets.
6. Underfitting and overfitting, handling mismatched classes (imbalanced datasets).
7. Performance metrics (precision, recall, F1, etc. And confusion matrix, cross validation, performance curves)

## Publications/ Conference Papers

Aug 2023      R. Edirisinghe, P. P. G. Dinesh Asanka, "Sentiment Reason Mining Framework for Analyzing Twitter Discourse on Critical Issues in US Healthcare Industry", INTERNATIONAL CONFERENCE ON SMART COMPUTING AND SYSTEMS ENGINEERING 2023, (<https://ieeexplore.ieee.org/document/10215010/>)

## Thesis & Projects

MSc Thesis - Twitter Sentiment Reason Mining Framework to Identify Major Problems in US Healthcare Industry. (Python based seamless NLP framework, Final Grade - A)

*This research presents a comprehensive NLP framework to study the US healthcare industry sentiment by applying machine learning and natural language processing techniques on Twitter data. The study aims to gain valuable insights into public sentiment towards healthcare-related issues and break down the overall twitter discourse on the USA healthcare domain. The framework presented involved multiple approaches including sentiment analysis, sentiment spike detection, clustering, keyword extraction, topic modeling, and textual association each contributing to a deeper understanding of the complexities within the healthcare domain. Social & Web Mining conducted through python web scraping algorithms emphasizing precision in mining targeted information. (snsraper, googlenews, beautifulsoup4). Multiple Sentiment Analysis models, Pre-trained: AFINN, TextBlob, and VADER, trained: Decision Trees, Logistic Regression, Random Forest, RNN\_TDFVec, CNN\_TDFVec, DNN\_TDFVec, AFINN, TextBlob, and VADER, evaluated with Accuracy, Precision, Recall, and F1 Score. Statistical Analysis and Spike Detection conducted with Z-scores & Matplotlib to detect sentiment spikes. Clustering and Vectorization approaches, K-Means, Agglomerative, Birch, MeanShift, CountVectorizer, TF-IDF, Word2Vec evaluated with Silhouette Score, Davies Bouldin Score, and the Calinski Harabasz Score. Topic Modeling approaches, Latent Dirichlet Allocation, Latent Semantic Analysis, and Non-Negative Matrix Factorization evaluated with Perplexity score, Coherence CV, Coherence U Mass, Coherence C UCI, Topic Diversity, and Topic Dominance Score. Textual Association and Similarity conducted through TF-IDF, Cosine Similarity, Sequence Similarity, and Levenshtein Distance.*

Undergraduate Thesis - Computer Aided Diagnosis System for Lung Cancer Detection (Standalone Java Based CADE system)

*This research aimed to develop a computer-aided diagnosis system for early detection of lung cancer using medical images. The system is designed to analyze x-ray and CT scan images to identify suspicious abnormalities indicative of lung cancer. Image processing techniques like segmentation, filtering, and feature extraction was utilized to detect lung nodules. The nodules were classified as benign or malignant using machine learning algorithms. The main goal was to reduce mortality rate from lung cancer by enabling early diagnosis improving treatment outcomes and survival rates. The project deliverables included system design, implementation, testing, user manual, and final software distribution.*

## Undergraduate Level Project -

*Student Accommodation System - Web based application using HTML, CSS, JavaScript, JSP, Servlet, and Google Map API.*

## HND Level Projects -

1. *Mugshot Identification System - SIFT based face identification from manipulated facial images based on Java SE, OpenCV, MySQL, & SIFT Algorithm.*
2. *Cab Reservation System - Web application based on .Net for online Cab Reservation based on ASP.NET, ADO.NET, Microsoft SQL, Google Maps, and JSP API.*
3. *Music Band Promotional Web site - Website based HTML5, CSS3, Bootstrap, JavaScript, jQuery, PHP, PDO, and MySQL*
4. *Client Management System - Web application based on JSP, Servlets, Microsoft Access*
5. *Employee Management System - Desktop Application based on Java SE, Microsoft Access*
6. *Student Grading System - Desktop Application based on Java SE, Microsoft Access*

## Core Competencies and Skills

### Soft Skills:

*Excellent communication and presentation skills (especially with non-technical stakeholders) | Strong tutoring and mentoring abilities | Proven expertise in business process improvement, design, and implementation | Strong analytical and problem-solving mindset | Ability to convey complex technical concepts in simple terms | Collaborative and proactive team player | Quick learner with a high adaptability to new tools and domains*

### Technical Skills:

*Machine Learning and Deep Learning | Natural Language Processing (NLP) and LLMs | Neural Networks and Transfer Learning | Statistical Analysis and Modeling | Text Reuse Detection and Semantic Change Analysis | Data Mining, Web Mining, and Social Mining | Predictive Analytics | Data Visualization and Dashboarding (Power BI, Tableau, Plotly) | Business Analysis and Workflow Automation | Collaborative Software Development and Version Control*

### Technologies:

Programming Languages: Python, SQL, Java, JavaScript, C#, R, VBA

LLMs & NLP: Hugging Face Transformers, spaCy, Gensim, NLTK, OpenAI API, Prodigy

Machine Learning & Deep Learning: PyTorch, TensorFlow, Keras, scikit-learn, XGBoost

Data Visualization & Analytics: Power BI, Tableau, QlikView, Plotly, Matplotlib, Seaborn

Cloud Platforms: Azure (Databricks, Synapse, Blob, Logic Apps), AWS (EC2, DynamoDB, Kinesis), MongoDB Atlas

Databases: MySQL, MongoDB, Amazon DynamoDB

Data Science Platforms: Weka, Orange

Web & App Development: HTML, CSS, ASP.NET, JSP

Automation & Scripting: VBA, Power Automate, Bash

## References

### [Dr. Rosane Minghim](#)

Professor - Artificial Intelligence, Data Analytics, & Algorithmics | Data Visualization and Human-centric Computing

Office 176 [School of Computer Science and Information Technology, Western Gateway Building, University College Cork](#)

Email: [r.minghim@cs.ucc.ie](mailto:r.minghim@cs.ucc.ie)  
Tel: +353 21 420 5901

### [Mushtaq Ahamed](#)

Director - Group Finance  
[Expolanka Holdings PLC](#)

15A, Clifford Avenue, Colombo 3, Sri Lanka

E-mail: [mushtaq@expolanka.com](mailto:mushtaq@expolanka.com)  
Tel: +94-77-7530355

### [Mushtaq Saheed](#)

Head of Finance - International,  
[EFL Global HQ](#)

10 Mile Post Ave,  
Colombo 00300.

E-mail: [mushtaq@efl.global](mailto:mushtaq@efl.global)  
Tel: +94-77-7 557226