Joji Yanagii
yanagiijoji@gmail.com

Summary
Data Scientist with 3 years of experience at Deloitte, specializing in transforming complex data into actionable insights for decision-making. Proficient in Python, SQL, Shell Script, Tableau, and machine learning, with proven success in leveraging advanced analytics to solve real-world problems. Skilled in cross-functional collaboration and stakeholder communication. Holds a Master's degree in Information Science.

Education
Master of Information Science
Nara Institute of Science and Technology (NAIST), 2021

Professional Experience
Data Scientist, Deloitte
Oct. 2021 - Present

- **Scientific Paper Analysis using Text Vectorization and Network Analysis (Education)**
  - Independently conducted in-depth text analysis and data visualization to assess the performance and interdisciplinary collaboration within research groups.
  - Leveraged Python, BERT models, clustering algorithms, and Tableau to extract insights, supporting strategic decision-making for academic evaluations.
  - Presented findings to clients, providing detailed reporting on research impact and collaboration networks.

- **Fraud Detection using Face Recognition and OCR on AWS (Government)**

    - Developed and deployed a fraud detection system using AWS Rekognition for image processing, integrating Face Recognition and OCR capabilities to identify counterfeit ID cards.
    - Successfully led the project, coordinating with experts to refine image processing accuracy, preventing over 100 cases of fraud and saving approximately 1 billion yen.
    - Tools: Python, Shell Script, AWS EC2, S3, and Linux.
- **Sentiment Analysis using LLM on Google Cloud for Marketing Decision Making (Retail)**
    - Conducted sentiment analysis on 3,000+ Twitter posts to determine public perception of products, using Vertex AI and Python for data collection, annotation, and LLM processing.
    - Delivered comprehensive sentiment reports, highlighting product reception trends and providing actionable insights for client marketing strategies.
- **Sales Forecasting for M&A Decision-Making using Machine Learning (Healthcare)**
    - Developed a sales forecasting model to aid M&A decision-making in the eldercare sector, with a focus on long-term revenue forecasting.
    - Employed scikit-learn to implement a linear regression model, enhancing accuracy through a logarithmic transformation of growth data, allowing clients to assess potential acquisition growth.
- **Online Survey to Estimate the Number of Baseball Fans (Sports)**
    - Led a targeted survey of 3,000+ respondents to estimate the size and demographics of the baseball fan base, supporting strategic marketing efforts.
    - Analyzed data with statistical techniques and clustering to identify key target demographics. Produced visualizations and PowerPoint presentations for client stakeholders.

Technical Skills

- **Programming Languages**: Python, SQL, Shell Script, C#
- **Data Science Techniques:**

    - Machine Learning (Clustering, Dimension Reduction, Regression Analysis)
    - NLP (Sentiment Analysis, Topic Modeling, Language Models)
    - Computer Vision (Image Processing, OCR)
    - Network Analysis (Centrality, Community Detection)
- **Tools & Platforms**: Tableau, Alteryx, ArcGIS, AWS (Rekognition, EC2, S3, SageMaker), Google Cloud (Vertex AI)
- **Statistical Analysis & Modeling**: Regression Analysis, Multiple Regression Analysis, Decision Tree, Hypothesis Testing and Significance Tests, PCA, Population Weighting, and Aggregation

---

Languages

- **Japanese**: Native
- **English**: Proficient
- **Arabic**: Basic/Conversational
