KEUN YOUNG YOON

|New York, Willing to Relocate| (934) 500-4949 | yoon.keunyoung.yky@gmail.com |
|youngyoon.me|www.linkedin.com/in/keunyoung-yoon | github.com/youngyoony

EDUCATION

State University of New York at Stony Brook *Expected 2026*
Master of Science in Data Science
Relevant Coursework: Data Structures and Algorithms, Data Analysis, Introduction to Probability

Korea National Open University, Seoul, Korea *2021*
Bachelor of Science in Data Science and Statistics, GPA: 3.64
Relevant Coursework: Data Mining, Unstructured Data Analytics, Data Processing and Applications, Multivariate Data Analysis, Forecasting Methods and Its Applications, Introduction to Statistical Deep Learning

Sogang University, Seoul, Korea *2015*
Bachelor of Business Administration and Political Science (Dual Majors)

SKILL

- **Analytics and Statistics** – Regression Analysis, Data Mining, Multivariate Data Analysis, Probability and Statistical Inference, Statistical Computing, Data Wrangling, ETL Processes, Forecasting Methods, Unstructured Data Analytics, Data Visualization
- **Big Data and ML/DevOps** – Apache Kafka, Apache Spark, Apache Airflow, Elasticsearch, RAG, LLM, LangChain, Vector DB, Docker
- **Data Visualization** – Tableau, Power BI, Matplotlib, Seaborn
- **Project Management**– Jira, Confluence
- **Programming Language and other** – Python, R, SQL, Pandas, PyCaret, Django, FastAPI, VBA, MS Office

EXPERIENCE

Nexon Korea *Gyeonggi-do, South Korea*
Senior Data Analyst, ELSWORD *2022 – 2024*

- Integrated Kafka and Spark to develop an advanced data streaming and real-time analytics pipeline, enabling efficient collection and preprocessing of in-game activity data and user behavior metrics. This integration improved data processing efficiency and reduced latency, allowing for immediate data-driven decision-making and enhancing the overall player experience
- Implemented a system that crawled external user community forums to analyze recommended specs for specific dungeons using Chroma and LangChain. This enabled us to provide item recommendation lists that matched user opinions, moving beyond simple log data to enhance the relevance of in-game recommendations.
- Conducted a detailed correlation analysis that disproved the hypothesis that owning more characters leads to higher spending, resulting in a shift in marketing strategies.
- Improved user retention metrics by analyzing D+30 retention and login frequency, resulting in a 12% increase in accuracy for measuring user engagement and improving overall retention rates.
- Analyzed monthly user attendance, identifying two key groups: short-term logins within 7 days and continuous logins for 28+ days. This insight drove a project to grow the upper-mid-range user segment by 20%.

Senior Data Analyst, FIFA Online 4 *2017 - 2022*

- Managed the microtransaction strategy and execution of 25+ series of paid in-game items, achieving the highest-ever revenue for the series by leveraging advanced data analysis and market trends.
- Analyzed complex probabilistic product purchase behaviors, discovering user patterns that improved user retention and drove a 10% year-over-year revenue growth.
- Validated VIP user behavior hypothesis, leading to a 12% increase in monthly revenue.
- Led the design and implementation of 20+ data dashboards, enhancing real-time analytics capabilities and enabling faster decision-making processes across the team.
- Achieved business outcomes through user behavior analysis: Analyzed non-core user behavior with free game currency and updated low-cost items with better rewards, increasing small-scale paid spending by 20%. Also, conducted an analysis of user responses, identifying patterns that led to adjustments in item configurations and a 15% increase in user satisfaction.

- Integrated various technical solutions such as re-integrating data protocols from TCP to UDP, developing an anomaly detection system, and creating a real-time notification service for probabilistic item data, resulting in a 20% reduction in system response times.
- Executed 8 long-term user satisfaction surveys, transitioning from basic trend analysis to segmented feedback, which provided actionable insights and led to a 25% improvement in tailored user experiences.
- Developed and implemented algorithms to track in-game economy metrics such as currency flow and spending trends, improving team decision-making efficiency.

Data Analyst, Need for Speed: Edge *2015 - 2017*

- Led the planning and execution of in-game monetization strategies and user retention events, contributing to a 25% increase in revenue and a 15% boost in user engagement.
- Conducted data-driven analysis of paid user behavior, identifying high-value players, which informed the adjustment of product offerings, resulting in a 20% increase in purchase frequency.
- Conducted core user preference analysis for vehicle performance in PC cafes, implementing a balanced update strategy that increased PC cafe playtime by 10% while maintaining appeal for home users.
- Planned and executed large-scale testing events, including three major Closed Beta Tests (CBTs) for a racing game with over 200,000 participants and one Focus Group Test (FGT) and Interview (FGI). Analyzed feedback from core users, improving user experience and engagement by evaluating both qualitative and quantitative data.

Data Analyst, FIFA Online 3 *2014 – 2015*

- Managed the execution of microtransaction and user retention events, driving a 15% increase in user engagement and a 10% boost in revenue within six months by leveraging paid in-game items and retention strategies.
- Analyzed and monitored in-game economy metrics, tracking production and deletion volumes of in-game currency, which improved market stability by 20%.
- Leveraged advanced analytical skills to interpret Key Performance Indicators (KPIs) such as Unique Users (UU), Churn Rate, and In-Game Play Metrics. These insights were used to refine daily, weekly, and monthly game performance reviews, contributing to a 12% increase in user retention and an 8% rise in average purchase per user.

RESEARCH | PROJECT

Creating a Document AI Chatbot Using Solar LLM and Kafka for Efficient Real-time Retrieval and Generation: Developed a Document AI chatbot using Solar LLM and Kafka for efficient real-time retrieval and generation, with a focus on building a RAG (Retrieval-Augmented Generation) chatbot designed to handle complex documents and provide real-time, accurate answers.

- Implemented a system to preprocess, split, and store PDF documents, ensuring efficient retrieval using Solar embeddings.
- Created a chatbot that can retrieve relevant document fragments and generate answers using LLM, providing real-time feedback to user queries.
- Used Kafka to manage real-time message flow, improving the scalability and response time of the chatbot.
- Reduced redundant document processing through caching, saving time and costs, especially for large files.

Streamlining Financial Data Processing: A Dockerized PySpark and Elasticsearch Pipeline for Real-time Data Visualization: Utilized financial API data to create an end-to-end processing pipeline with Docker, PySpark, Elasticsearch, and Kibana for real-time financial data visualization.

- Deployed a PySpark web server using Docker Compose to ensure seamless execution and management of the pipeline.
- Verified the functionality of the PySpark server through Jupyter notebooks, confirming smooth data processing.
- Integrated Elasticsearch to store the processed data, ensuring accurate and efficient data storage and retrieval.
- Configured a DAG (Directed Acyclic Graph) in Airflow to automate and schedule the data processing tasks, ensuring timely updates and robust pipeline management.
- Visualized the financial data using Python, providing insightful dashboards that support data-driven decision-making

MMORPG Player Behavior Prediction and Analysis: Analyzed large-scale MMORPG player data using Python, Pandas, Seaborn, and Matplotlib to uncover relationships between combat power and in-game activities. Employed PyCaret, Gradient Boosting Regressor, and MiniBatch KMeans for predictions and segmentation, focusing on user engagement and retention strategies.

- Utilized Python, Pandas, Seaborn, and Matplotlib to analyze large-scale MMORPG player data, focusing on combat power progression and in-game activity engagement (e.g., dungeons, quests).
- Applied PyCaret, Gradient Boosting Regressor, and Python to predict login days based on game data. Additionally, used these tools to classify future combat power levels, identifying patterns in player engagement.
- Utilized MiniBatch KMeans and Python for user segmentation to support retention strategies.