NLP: IAB Content Classification

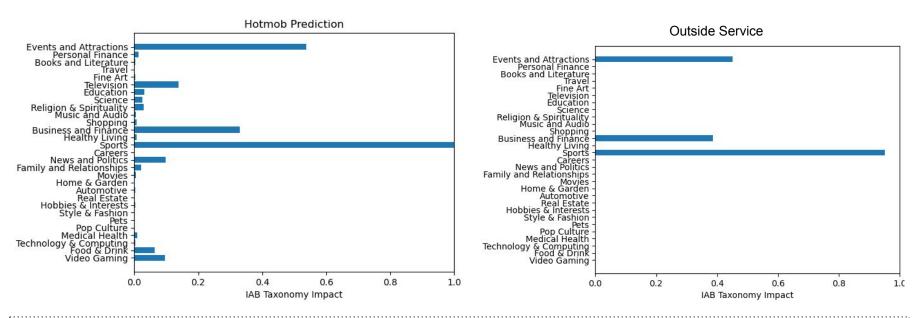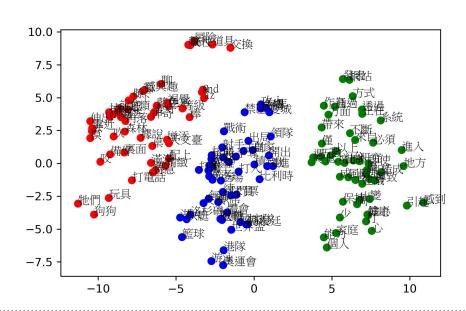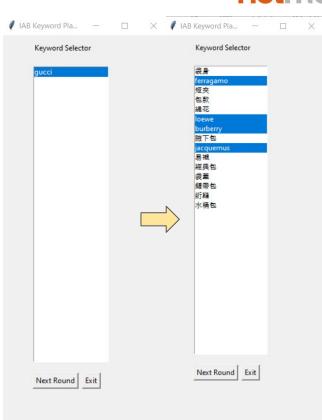Hotmob Prediction — Outside Service

- My Classifier demonstrates a **strong alignment of 89%** with an external service, which serves as the training + ground truth.
- Moreover, it operates at a fraction of the cost, running efficiently at **just 1/20th of the expense**.
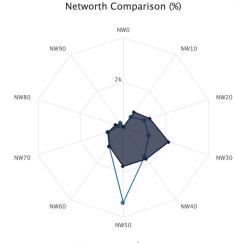
NLP: Internal-Use Tools

Designed and implemented advanced tools for extracting synonymous terms and assessing audience qualifications, resulting in a significant reduction of time spent per targeting campaign **from three days to just two hours**.

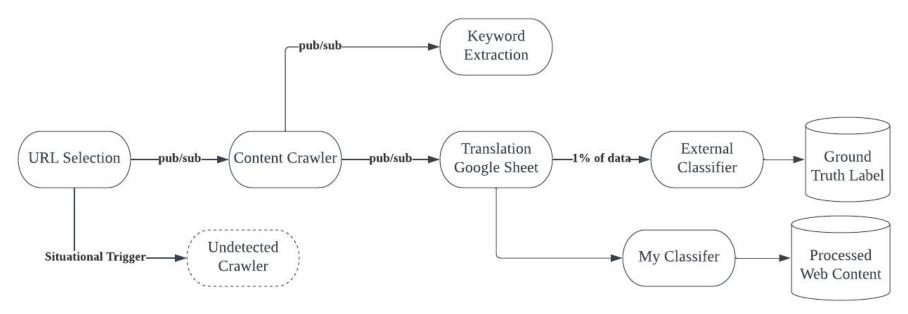# NLP: Streamlined Visualization

Implemented streamlined workflow, reducing creation time of visuals for pitching decks, insight reports, and social media posts **from 3-7 days to 1 day.**
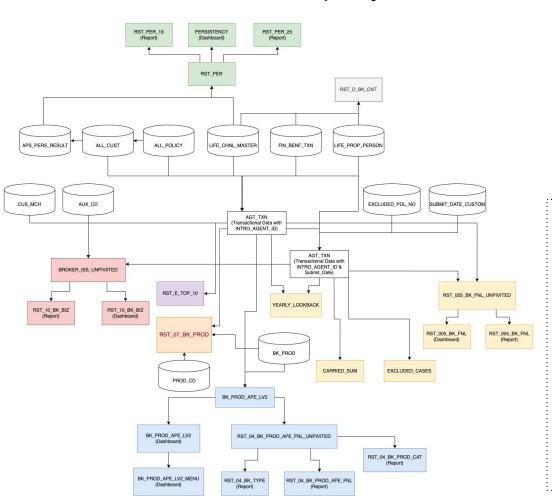
Web Content Pipeline

Implemented a comprehensive and **fully automated web crawling system**, enhanced with the utilization of Serverless Computing and Pub/Sub on the Google Cloud platform. This transformation significantly improved stability compared to the previous manual operation, which was prone to frequent failures.

Global Insurance Company: Internal Data Modelling Revamp
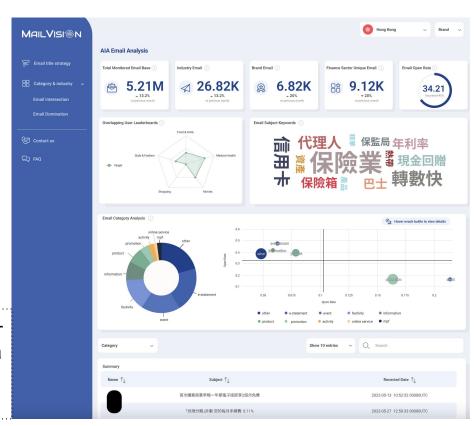

- Revamped a Global Insurance Company's data modeling system from SAS to Power Query
- Demonstrated strong communication skills, provided professional suggestions, and delivered clear visual explanations that **impressed their Business Intelligence Team**
- Cultivated and maintained a long-lasting client relationship through exceptional performance, resulting in ongoing collaboration to date
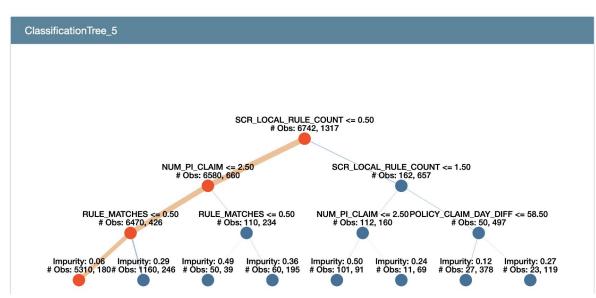
User-Focused Dashboards

Led production of user-focused dashboards for internet activity and mailbox usage. Transformed data delivery to **prioritize actionable insights** by integrating multiple comparison figures

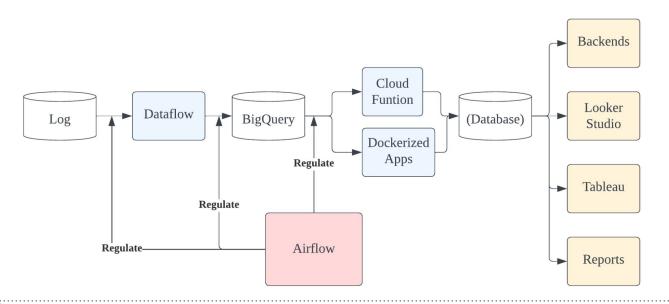# Global Insurance Company: Machine Learning Exploration

hotmob

Select Component to Graph

0 : Predicted (0) | PrScore (0.967) | Actual (NaN)

ClassificationTree_5

SCR_LOCAL_RULE_COUNT <= 0.50
Obs: 6742, 1317

NUM_PI_CLAIM <= 2.50
Obs: 6580, 660

SCR_LOCAL_RULE_COUNT <= 1.50
Obs: 162, 657

RULE_MATCHES <= 0.50
Obs: 6470, 426

RULE_MATCHES <= 0.50
Obs: 110, 234

NUM_PI_CLAIM <= 2.50
Obs: 112, 160

POLICY_CLAIM_DAY_DIFF <= 58.50
Obs: 50, 497

Impurity: 0.06
Obs: 5310, 180

Impurity: 0.29
Obs: 1160, 246

Impurity: 0.49
Obs: 50, 39

Impurity: 0.36
Obs: 60, 195

Impurity: 0.50
Obs: 101, 91

Impurity: 0.24
Obs: 11, 69

Impurity: 0.12
Obs: 27, 378

Impurity: 0.27
Obs: 23, 119

- Developed, hosted, and optimized Azure fraud detection model
- Explored **ML explainability** for inference enhancements
- Impressed client through **exceptional out-of-scope exploration and preparation**

Data Preprocessing

- Collaborated with technical team to fully automate end-to-end data processing pipeline using Airflow
- Identified and resolved issues with the original ID system, improving **efficiency and scalability by 14x**
- Implemented more efficient algorithms, resulting in a **50% reduction in processing time**
- Enabled seamless generation of log files and client-facing reports, enhancing overall productivity and efficiency