LAGUNA STATE POLYTECHNIC UNIVERISTY -
SANTA CRUZ CAMPUS

DIGI-BAYANI SYSTEM:

CLASSIFYING FILIPINO & TAGLISH SOCIAL MEDIA POSTS FOR URGENT DISASTER RESPONSE

BSCS 3B
Machine Learning (CSST101) &
KRR (CSST102)

Bulos, Mariabel M.
Ramos, Railey Mar M.
Rosario, Keanne Jericho M.

PROBLEM & MOTIVATION

Gap: Information overload on social media during typhoons.

Language Barrier: High usage of Taglish and code-switching makes standard English classifiers fail.

Stakes: Delay in identifying "Urgent Need" leads to slower response times.

SYSTEM OVERVIEW (THE HYBRID APPROACH)

LAYER 1 (ML)

Statistical classification of intent.

LAYER 2 (KRR)

Rule-based reasoning for severity and resource allocation.

SDG ALIGNMENT

Supporting SDG 11 (Sustainable Cities) and SDG 13 (Climate Action).

THE MACHINE LEARNING PIPELINE

FEATURE EXTRACTION

TF-IDF Vectorization

DATASETS

Combined 11,212 records from Typhoon Yolanda and 2012 Floods.

PREPROCESSING:

Regex-based cleaning (stripping URLs, @mentions, and hashtags).

ALGORITHM

Logistic Regression (Liblinear Solver)

PERFORMANCE RESULTS

Overall Accuracy: 77.80%

Key Strength: 98% Recall for 'Disaster-Related' class.

Insight: The model is optimized to minimize "false negatives"—we would rather investigate a minor post than miss a major emergency.

KNOWLEDGE REPRESENTATION & LOGIC

SYSTEM

Stack: Python (Flask), HTML5, CSS3, JavaScript.

Interface: Clean, dark-mode dashboard for 24/7 monitoring.

Process: User pastes tweet → Async API Call → Real-time Label & Severity output.

CONCLUSION

Success: Developed a functional Taglish-aware classifier.

Next Steps:

- Implementing SMOTE to balance minority classes.
- Integration with Geolocation APIs for map-based tracking.
- Deep Learning (BERT/LSTM) for better syntax understanding.

REFERENCES

- CrisisLex.org (Primary Dataset).
- Scikit-Learn & Flask Documentation.

THANK YOU!