MEMBERS ASSIGNMENTS REPORT

Group Name:

	Member 1	Member 2
Name	Keilor Fallas Prado	Ky Dang
Email	kfallasprado@gmail.com	Keith.dang1610@gmail.com
Country	Costa Rica	Vietnam
College/ Company		
Specialization	NLP	NLP

Division:

Weeks	Plan	Ky Dang	Keilor
Week 8	Review data source and ensure it is representative of hate speech contexts.	Data Review	Data Review
Week 9	Remove duplicates, nulls, and irrelevant data.	Data Preprocessing	Data Preprocessing
Week 10	Evaluate and select models such as Logistic Regression, SVM, or Transformers (e.g., BERT).	Using emojiBERT (ver1)	Using BERT-based
Week 11	Tokenization- Identify relevant linguistic and contextual features	Manual Tokenizer utilizing DistilBERT (ver2)	Using directly BERT tokenizer
Week 12	Training and evaluation model	Used DistilBERT (ver2) and run model.	Rerun and optimize distilBERT model (previous version)
Week 13 (optional)	Rerun improved model	Rerun the improved optimize distilBERT model (final version) with accuracy of detection hate class	Rerun the improved optimize distilBERT model (final version) with accuracy of detection hate class

Deliveries

ID	Item	Url
1	Compressed model (.zip)	model_link
2	Notebook file (.ipynb)	Notebook_Link