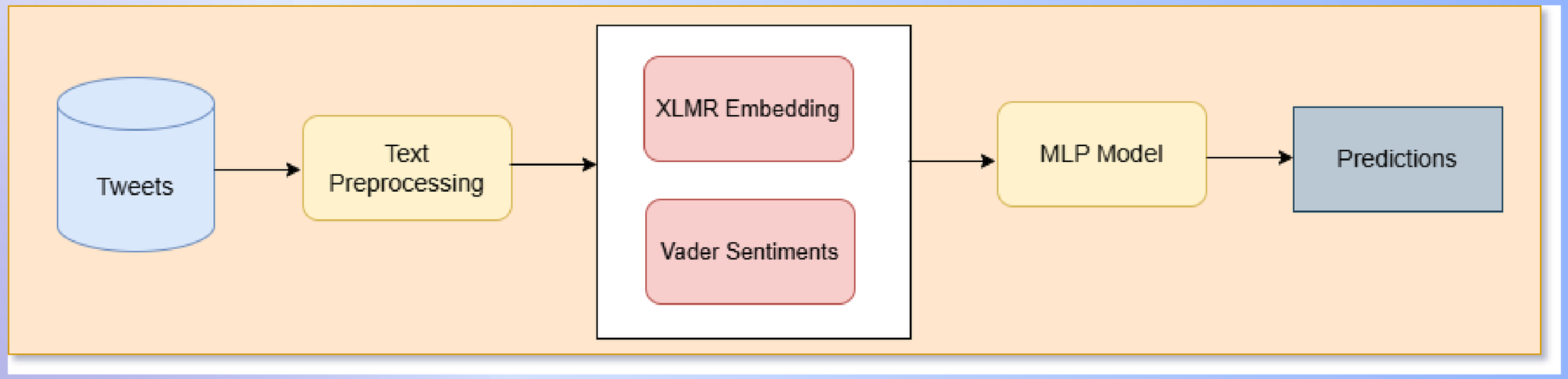
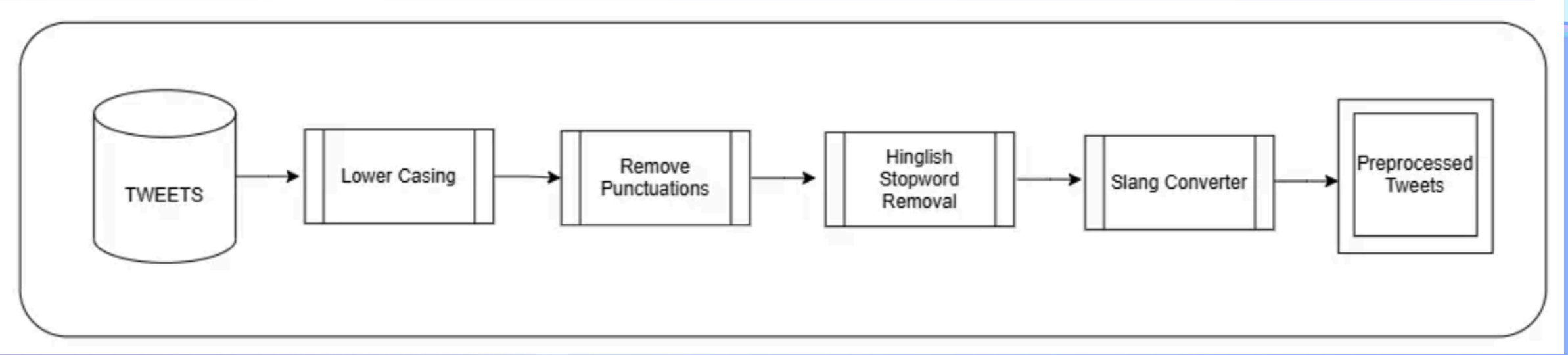
MULTILIGULAL SARCASM DETECTOR

TEAM - DATA MAVERICKS

APPROACH FLOW DIAGRAM



PREPROCESSING STEPS



SENTENCE EMBEDDING TECHNIQUE

Key Features of XLM-R Embeddings

- **Contextual Representations: Captures the meaning of words based on surrounding context.**
- **Multilingual Capability: Supports 100+ languages without relying on parallel data.**
- **Robust Performance: Outperforms previous multilingual models on several NLP benchmarks.**

SENTIMENT ANALYSIS USING VADER

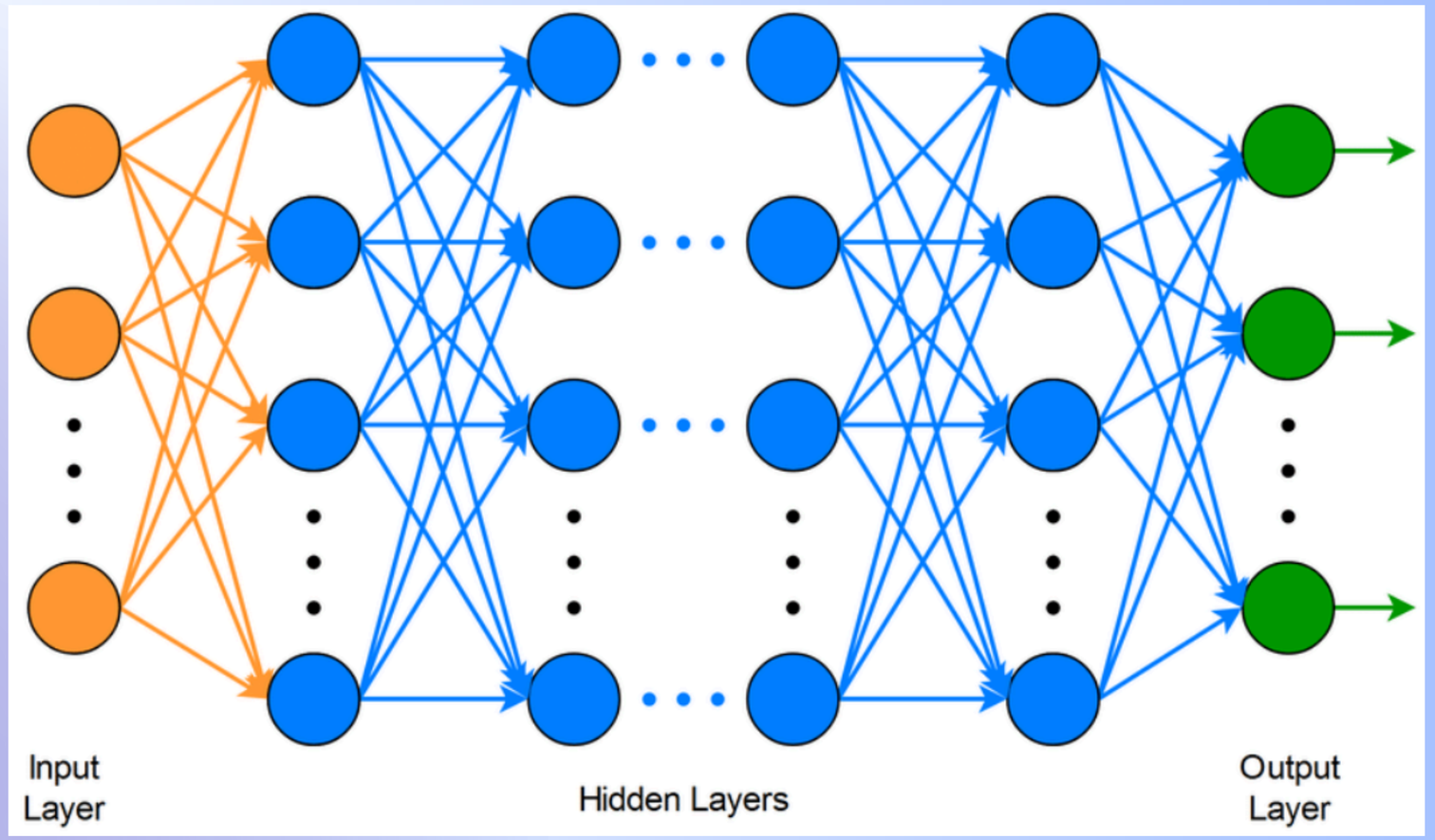
How VADER Works?

- Uses a predefined lexicon where words are assigned sentiment scores.
- Produces a compound score between -1 (negative) and 1 (positive):
- Positive Sentiment: Compound score > 0.05
- Neutral Sentiment: Compound score between -0.05 and 0.05
- Negative Sentiment: Compound score < -0.05

Why VADER?

- Works well on short, social media text.
- No need for large training datasets (lexicon-based approach).
- Captures intensity modifiers (e.g., "very good" is more positive than "good").

MULTILAYER PERCEPTRON MODEL



https://www.researchgate.net/figure/The-generic-structure-of-MLP_fig2_355118998

HYPERPARAMETER	DESCRIPTION	VALUES
Input Layer	Number of input features	772
Hidden Layers	Number of hidden layers and nodes in each layer	[512, 256, 128, 64]
Output Layer	Number of output nodes and activation function	1 node, Sigmoid activation
Activation Functions	Activation functions for hidden and output layers	ReLU (hidden), Sigmoid (output)
Loss Function	Loss function for binary classification	Binary Crossentropy
Optimizer	Optimization algorithm	Adam
Learning Rate	Learning rate for the optimizer	0.0001
Batch Size	Number of samples per gradient update	32
Epochs	Number of training iterations over the entire dataset	100
Early Stopping	Early stopping criteria to halt training if validation loss doesn't improve	Patience = 10, Monitor = val_loss
Metrics	Metrics to evaluate model performance	Accuracy, Precision, Recall

EMBEDDING	MODEL	Acc	Macro F1	Weighted F1
FAST TEXT + Vader Sentiments	XGBoost	0.95	0.94	0.95
	LightGBM	0.94	0.93	0.94
	AdaBoost	0.93	0.92	0.92
	MLP	0.96	0.95	0.96
XLmr + Vader Sentiments	XGBoost	0.97	0.96	0.96
	LightGBM	0.97	0.97	0.97
	AdaBoost	0.96	0.96	0.96
	MLP	**0.98**	**0.98**	**0.98**
mBert + Vader Sentiments	XGBoost	0.97	0.96	0.97
	LightGBM	0.97	0.96	0.97
	AdaBoost	0.96	0.96	0.96
	MLP	0.97	0.97	0.97
XLmr (without Vader Sentiments)	XGBoost	0.96	0.95	0.96
	LightGBM	0.95	0.95	0.95
	AdaBoost	0.95	0.95	0.95
	MLP	**0.97**	**0.97**	**0.97**

THANK YOU!