EMOTION RECOGNITION FROM TEXT

KENDRICK HUANG
NIM : **2702341606**

ALLENANDA SHEVA ATHALLAH
NIM : **2702312114**

Latar Belakang

Twitter sering digunakan untuk menyampaikan opini, kritik, dan diskusi. Namun, banyak pengguna memakai sarkasme atau ironi, yang maknanya tidak selalu sesuai dengan kata-kata literal. Hal ini menyulitkan sistem analisis otomatis. Karena itu, dibutuhkan model yang mampu mengenali tweet sarkastik, ironis, atau netral.

DATASET

https://www.kaggle.com/datasets/nikhiljohnk/tweets-with-sarcasm-and-irony

Dataset ini terdiri dari 81.407 tweet untuk train dan data untuk di test sebanyak 8.127 yang masing masing memiliki class, dengan pembagian data train dan test menggunakan rasio 80:20.

Dataset ini memiliki 2 kolom :

- tweets
- class

Class:

- Figurative
- Sarcasm
- Irony
- Regular

PREPROCESSING

01.

LOWERCASING

02.

REMOVING

- URLs
- #hashtags
- @mentions
- excess whitespace

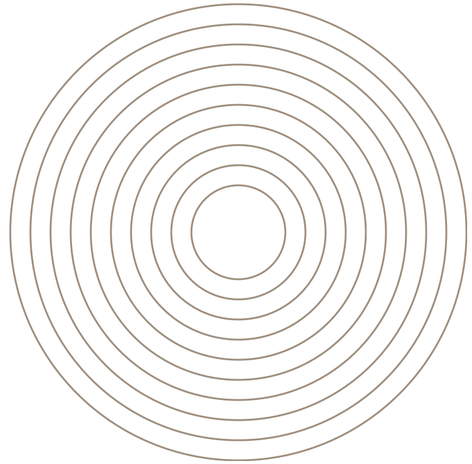
03.

TOKENIZATION
(KERAS TOKENIZER)

04.

PADDING
(MAX_LEN = 40)

Model Architecture



Input Text
↓
Tokenizer
↓
Embedding Layer
↓
SpatialDropout1D
↓
BiLSTM (context understanding)
↓
CNN (n-gram sarcasm trigger extraction)
↓
GlobalMaxPooling1D
↓
Dense(128) + Dropout
↓
Softmax Output (class probabilities)

Why Hybrid BiLSTM + CNN

BILSTM

- captures sarcasm context both forward & backward

CNN

- detects local irony triggers (exaggeration, slang sarcasm)

TOGETHER

- balanced global context + micro sarcasm patterns

Training Strategy

01. LOSS: LOSS: SPARSE CATEGORICAL CROSSENTROPY

02. OPTIMIZER: ADAM (LR = 2E-3 → AUTO REDUCED)

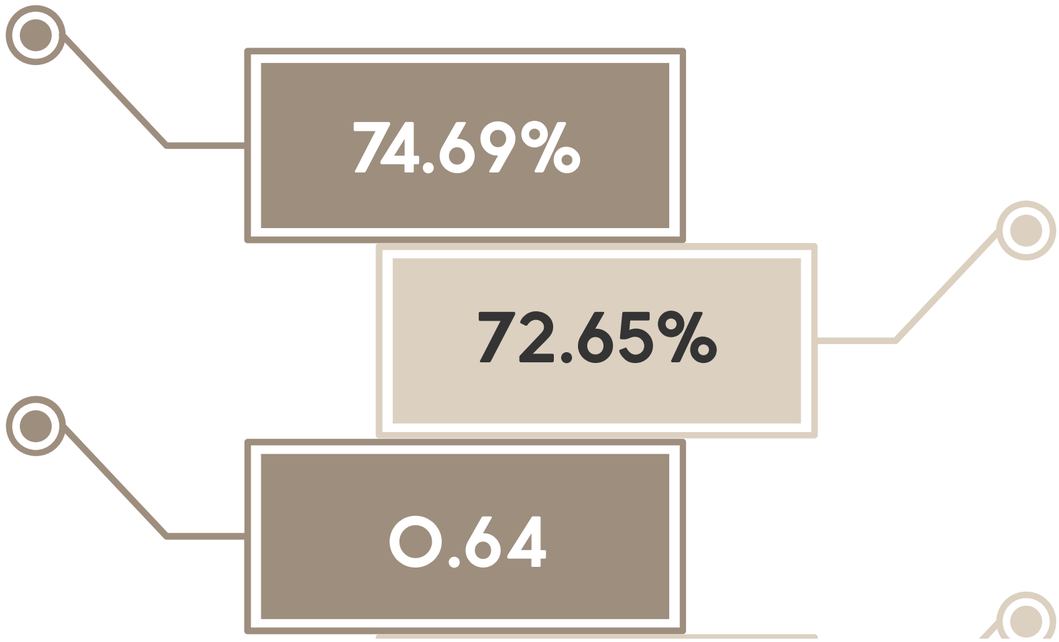
03. EARLYSTOPPING (PATIENCE = 4)

04. REDUCELRONPLATEAU (FACTOR = 0.5)

05. MODELCHECKPOINT (BEST WEIGHTS SAVED)

Performance Results

Training Accuracy



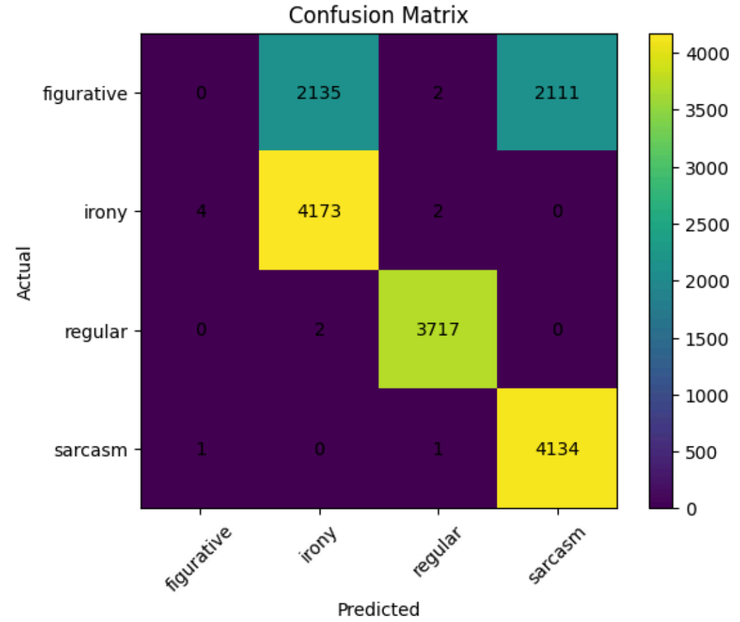
Validation Accuracy

F1 Score

Confusion Matrix

Classification report:

	precision	recall	f1-score	support
figurative	0.00	0.00	0.00	4248
irony	0.66	1.00	0.80	4179
regular	1.00	1.00	1.00	3719
sarcasm	0.66	1.00	0.80	4136
accuracy			0.74	16282
macro avg	0.58	0.75	0.65	16282
weighted avg	0.57	0.74	0.63	16282



Brief explanation:

- Misclassification mostly happens between sarcasm → regular
- Indicates difficulty detecting implicit tone without contextual cues

Demo

Manual Input Example

“Test on a Saturday! Thank you uni! #sarcasm @ Griffith University, Nathan Campus”	**Sarcasm**
“If it's possible for my street to be louder today while I'm on calls during WFH, that would be great #sarcasm”	**Sarcasm**
“Too good not to share! #Gerrymandering #voting #irony http://t.co/UmvVptBHpU”	Irony

Reflection

- Model understands sarcasm patterns, but:
  - lacks pragmatic world knowledge
  - cannot detect culturally dependent humor
- Sarcasm ≠ purely textual → needs external context

Improvement Ideas

- Add Transformers (BERT / RoBERTa)
- Emoji sentiment fusion
- Multitask sarcasm + sentiment + emotion
- Add speaker-level context metadata

Conclusion

- Hybrid BiLSTM–CNN effectively learns sarcasm expressions
- Strong classification for explicit sarcasm
- Overfitting suggests need for deeper contextual modeling
- Demonstrates feasibility for moderation & social media interpretation tasks

Thank You