

# Small Dose, Big Effect: Leveraging Synthetic Data for Candy Speech Detection

Sophie Luise Opheiden, Elisei Shushpanov, Faeze Ghorbanapour

## The "What & Why"

### - Motivation

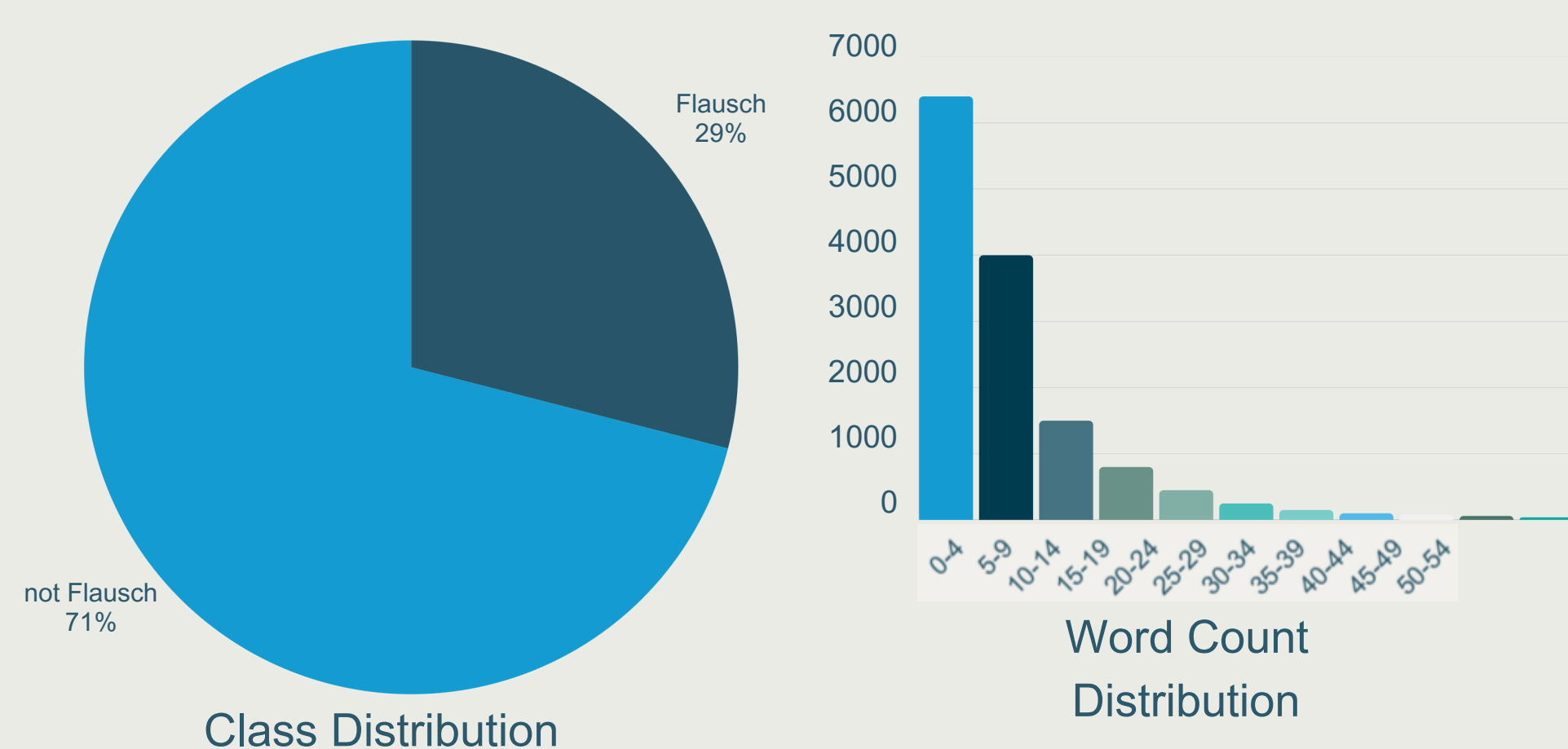
- While most research focuses on censoring negative speech, the active promotion of positive online discourse is often overlooked.
- Our goal is to detect "Candy Speech" to help foster more supportive and empathetic online communities.

## What is 'Candy Speech'?

- Definition:** An expression of positive attitudes on social media toward individuals or their output (videos, comments, etc.). It's the opposite of hate speech. **Examples:**
- ✓ "Weiter so daumen hoch :)" (Keep it up, thumbs up :))
- ✓ "Ihr seid einfach der Hammer!" (You guys are just awesome!)
- ✓ "Echt geiles Lied ❤️" (Really cool song ❤️)

## Dataset

- Source:** German-language YouTube comments
- Key Challenge:** The dataset is highly imbalanced, with informal language and short texts.



## Key Insight

- 50% of comments have 6 words or fewer.
- "Noisy" text (emojis, slang, typos) is a critical predictive signal, not just noise to be removed.

## GitHub Repository



## Contact

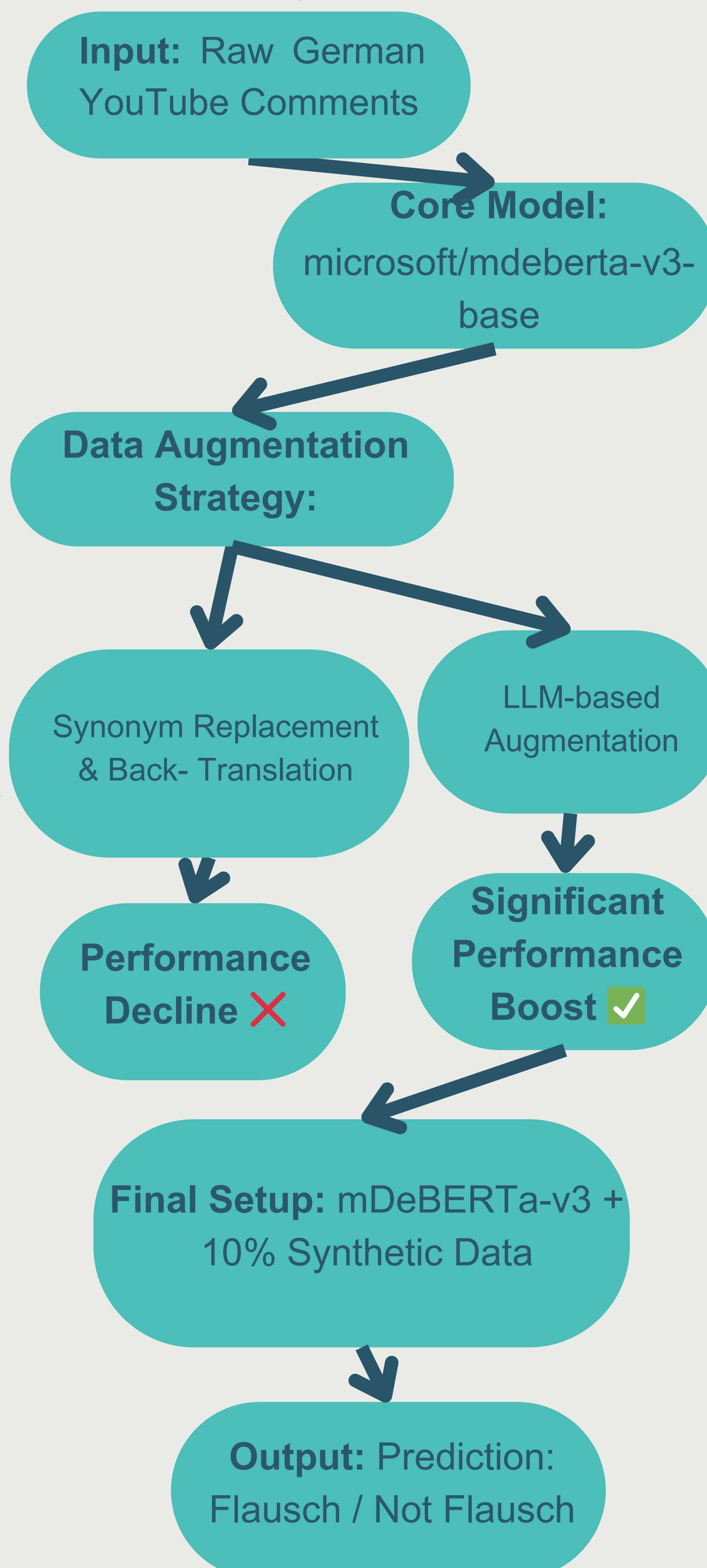
- sophie.opheiden@tum.de
- e.shushpanov@tum.de
- faeze.ghorbanpou@tum.der

## The "How"

### - Methods Explored

- Data Preprocessing:**
  - Initial tests with aggressive cleaning (lemmatization, etc.) hurt performance.
  - Final models used raw text.
- Model Architectures:**
  - Classical Baselines: Logistic Regression, SVM, etc. (with TF-IDF, Word2Vec).
  - Full Fine-Tuning: mBERT, XLM-RoBERTa, mDeBERTa (Best).
  - Parameter-Efficient Fine-Tuning (PEFT): LoRA & Prompt Tuning were tested as computationally efficient alternatives.

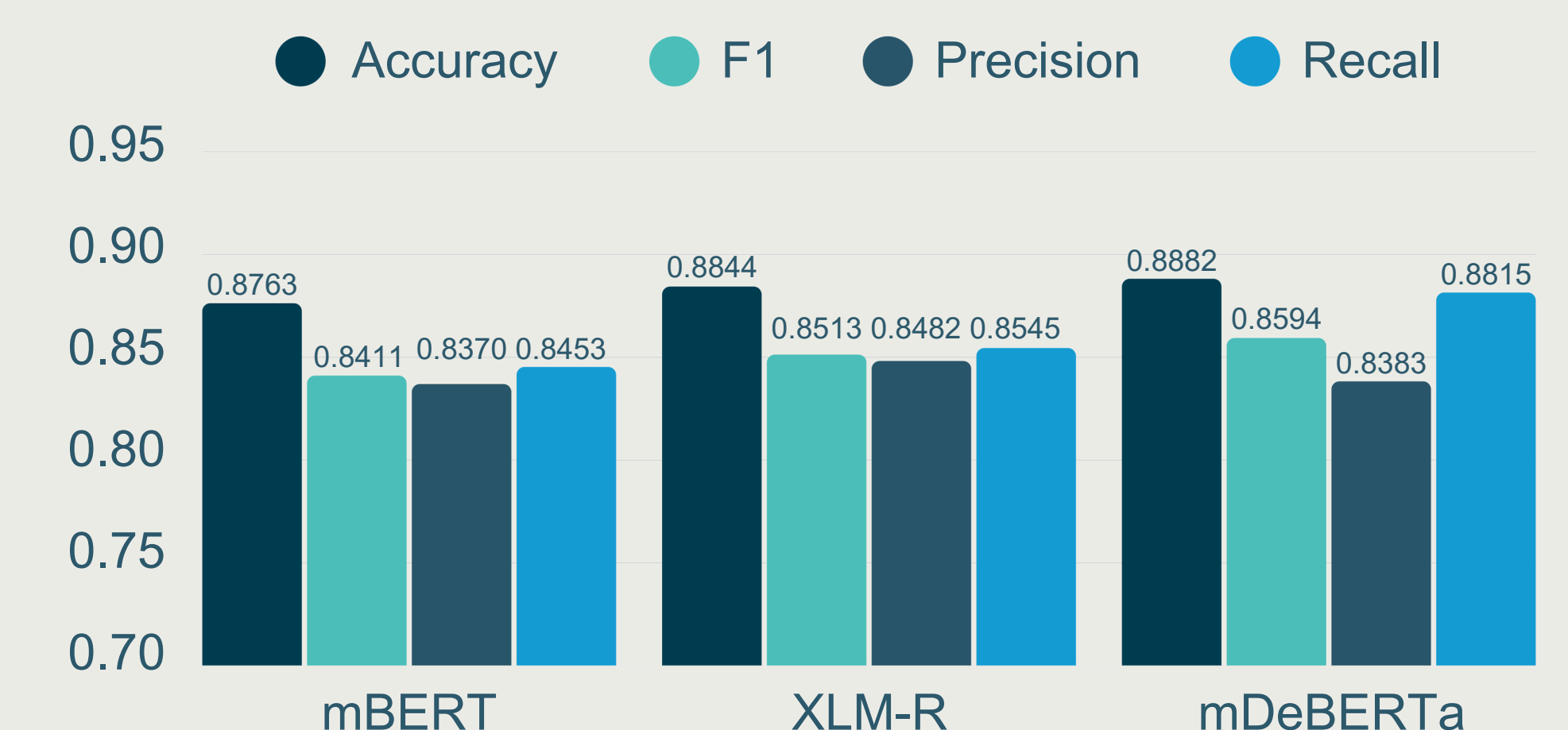
## The Winning Pipeline



## The "So What?"

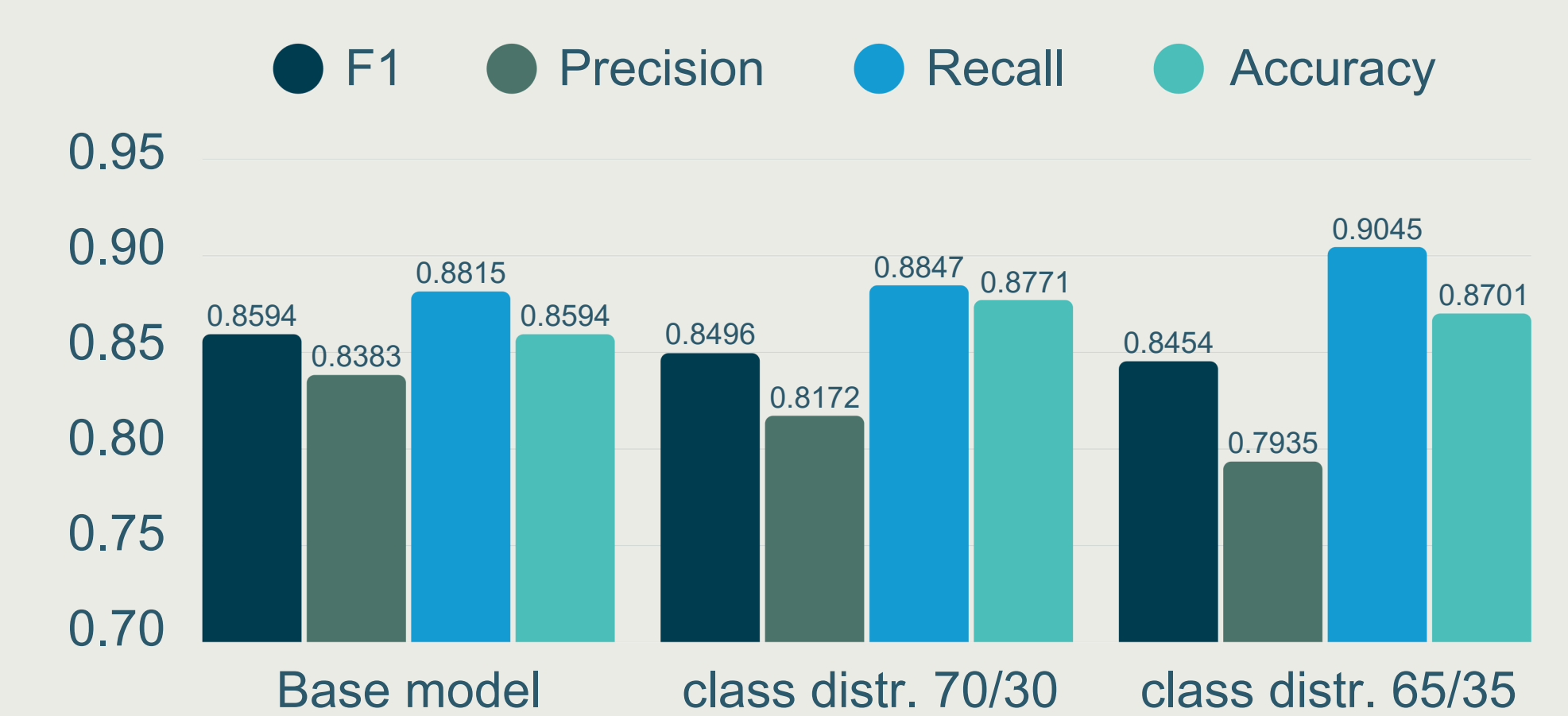
### - Model Comparison

#### mDeBERTa Outperforms Other Transformers



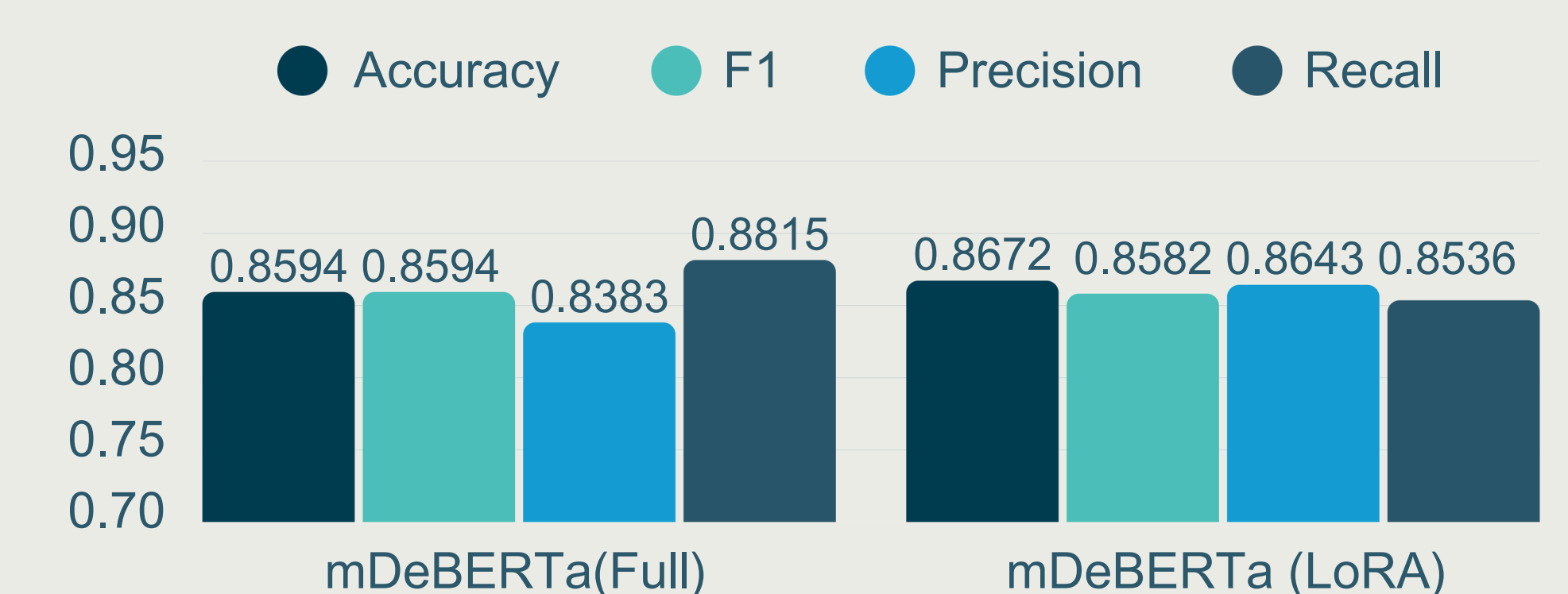
## Augmentation Sweet Spot

### Small Dose, Big Effect (mDeBERTa)



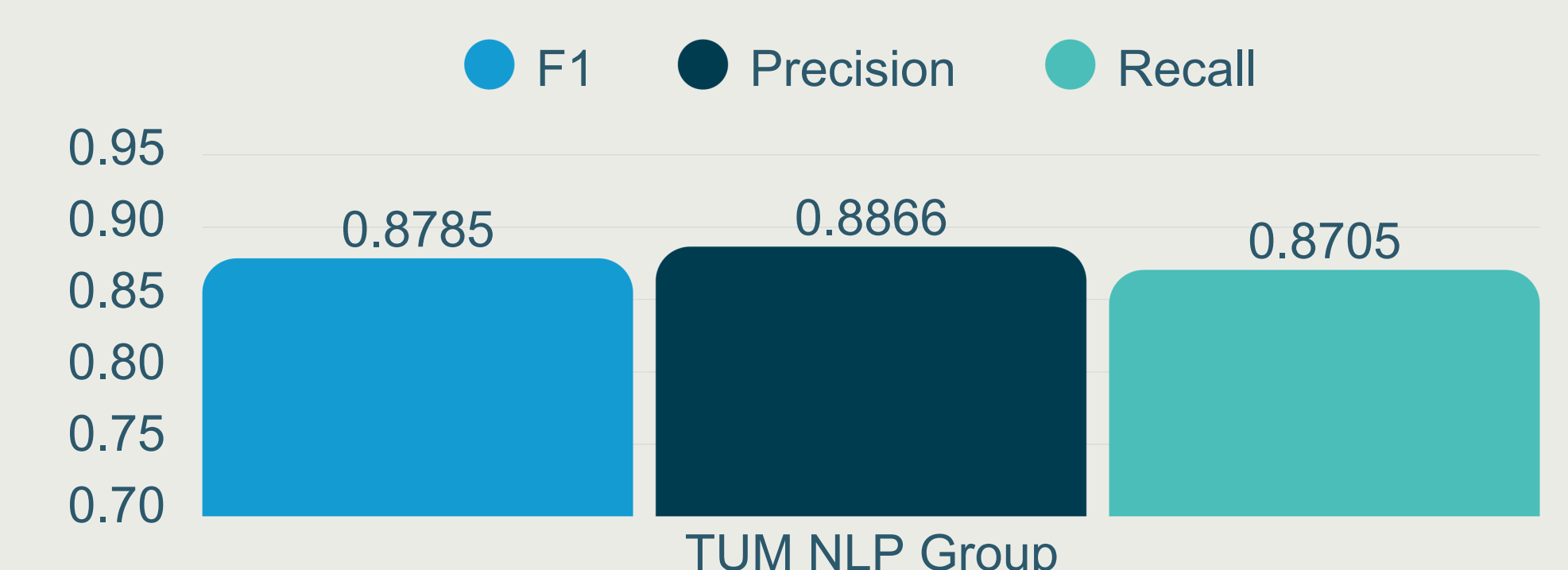
## PEFT vs. Full Fine-Tuning

### LoRA is an Efficient Alternative



## Official Result

### Achieved Rating: 6th / 20



## Takeaways & Conclusion

- Data > Model:** A 10% data augmentation with an LLM gave a bigger boost than changing model architecture.
- Efficiency Matters:** PEFT (LoRA) achieves results comparable to full fine-tuning at a fraction of the computational cost.
- "Noise" is Signal:** Informal language ("noisy" text) is a key feature, not a bug to be removed.
- Future Work:** The most promising direction is expanding the high-quality dataset.