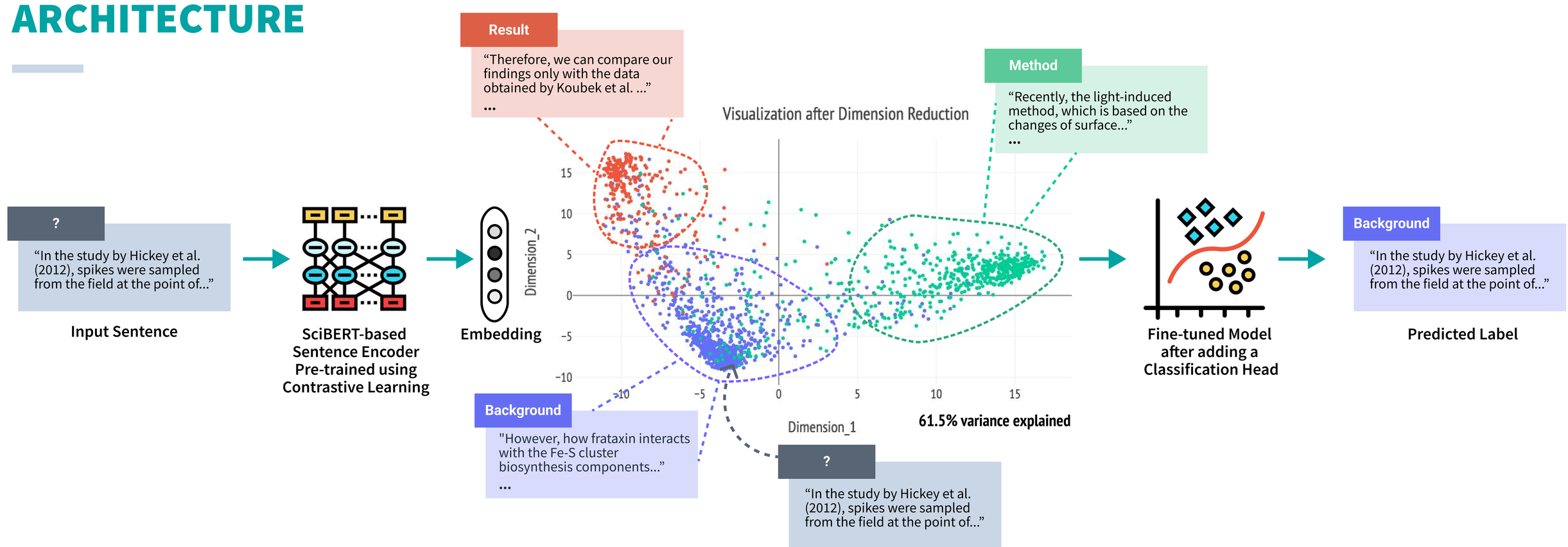
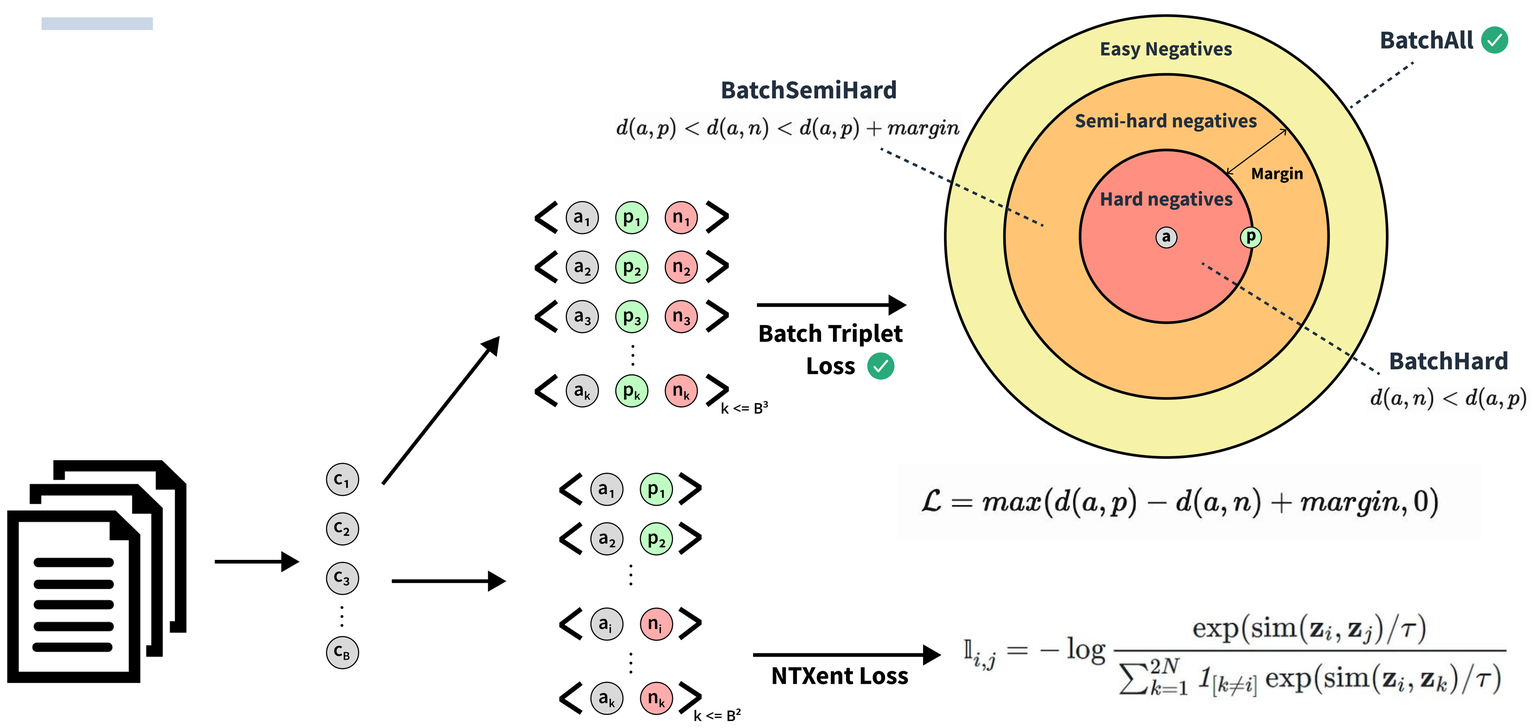
Constrastive Pre-training for Citation Intent Classification

Goh Jun How, Tan Xing Rong James, Motwani Richa Vijay, Pratham Jain, Sean Lam, Anjali Agarwal

ARCHITECTURE



CONTRASTIVE PRETRAINING

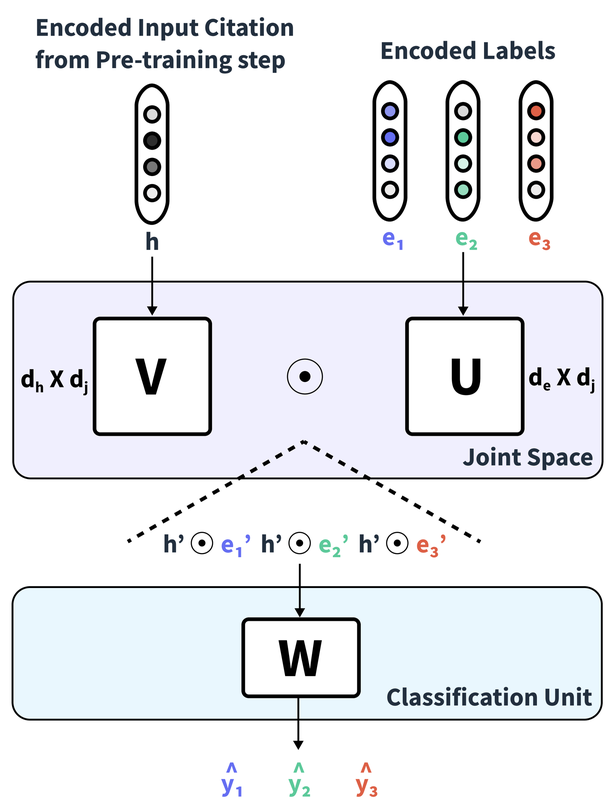


1. Take a batch of **B** samples and compute their embeddings
2. Find all valid triplets / pairs **<anchor, +ve, -ve> / <anchor, +ve> & <anchor, -ve>**
3. Train SciBERT-based model to bring positive samples closer in the embedding space and negative samples further apart

Supervised Approach — Using labels
< a, +ve > same intent
< a, -ve > different intent

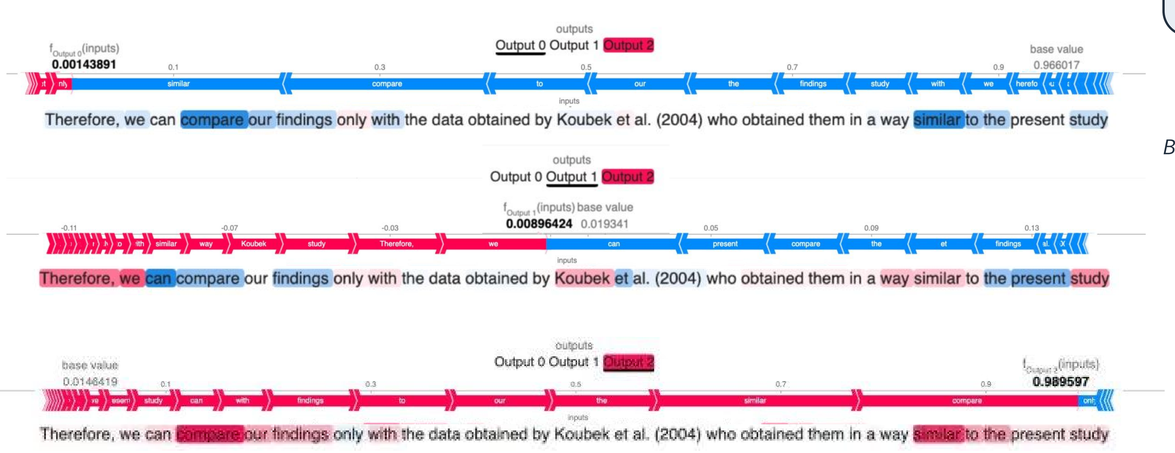
Self-Supervised Approach — Using section names
Section Name Embedding of
< a, +ve > high cosine similarity
< a, -ve > low cosine simiarity

FINETUNING WITH JOINT INPUT-LABEL EMBEDDING



INTERPRETABILITY & TRANSPARENCY

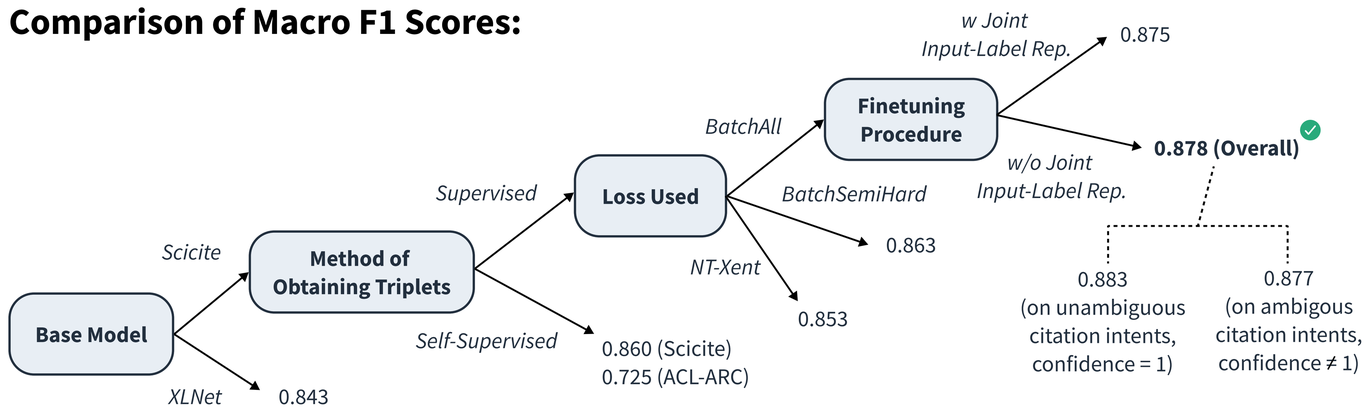
Text_instance: "Therefore, we can compare our findings only with the data obtained by Koubek et al. (2004) who obtained them in a way similar to the present study."



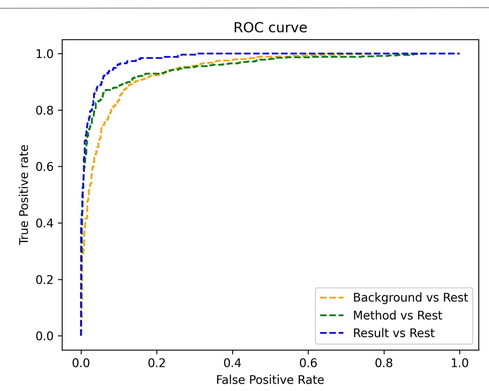
RESULTS

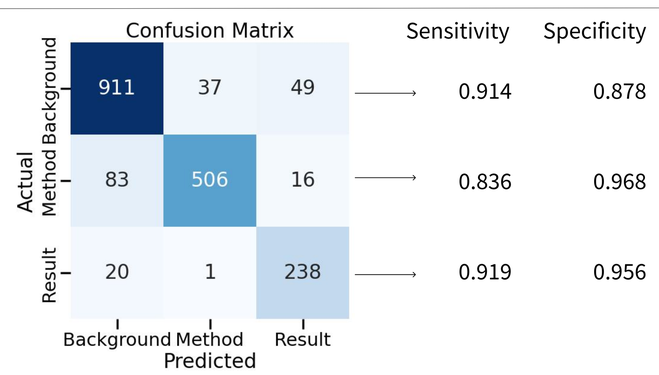
Contrastive pre-training creates embedding clusters that improves classification performance

Comparison of Macro F1 Scores:



Baseline Model (w/o Contrastive Pretraining): 0.855





SciBert w BatchAllTripletLoss (sup.)

Project ID: CS4248-10

Sponsors: Micron, Goodnotes, sea, Optiver, SingHealth