	Vector Embeddings			
	(NLP)			(Ethics)
Vector Elements	Title	Post Content	Hashtags	Metadata x15 features
Data Type	Separate Raw Texts			Categorical or Numeric Data
Methods	Text Normalize			Encoding Data: One-hot / LabelEncoder
	Vectorize	Vectorize	Vectorize	
	(Optional) Hyper-Parameter Tuning			
	Baseline			Any kind of simple ML model (e.g. scikit-learn decision trees, random forest, SVM, …) One or more is also fine.
	Deep Learning Fine-Tuning			
	LLM-based method			
	Hybrid: Best 2 models above			Control Experiment: Full Dataset vs non-sensitive attributes only
	Cross-Validation			
Goal	**Label which posts are suicidal** (binary classification)			Using NLP results as annotations,
				Annotated Labels
	Val Set	evaluation metrics		Find whichever features (subgroups) make impact on text predictions
	Test Set	no correct labels, annotate in Result.jsonl		
				Explanability: SHAP / LIME Model