TBD End-to-End Workflow: Subject to change.

Workflow Stage	Technology / Tool	Role in Project	Required / Suggested
1. Data Collection	–	Gather forum/social/clinical text, labels	Required
2. Preprocessing & Tokenization	Hugging Face Tokenizers	Convert raw text → input IDs, masks; handle subwords, padding, OOV	Required
3. Embedding Layer	Word2Vec	Unsupervised static embeddings on large unlabeled corpora	Suggested
	ELMo	Deep contextual LSTM embeddings based on sentence context	Suggested
	BERT	State-of-the-art bidirectional Transformer embeddings	Required
4. Model Construction	Keras (TF backend)	Build classification head on top of embeddings (Dense + activation + dropout)	Required
5. Training & Fine-Tuning	TensorFlow Runtime	Autodiff, GPU/TPU acceleration, mixed precision	Required
	Hugging Face Trainer	Ready-made training loop, batch management, checkpointing	Suggested
6. Evaluation	Scikit-learn / TF metrics	Focus on **Recall** (Sensitivity) to minimize false negatives	Required
7. Interpretability	LIME	Local surrogate models for per-instance explanations	Suggested
	SHAP	Global and local attributions via Shapley values	Suggested
8. Deployment	TensorFlow Serving	Serve your Keras/TF model behind a scalable API	Required
	Hugging Face Inference API	Cloud-hosted endpoints for your fine-tuned model	Suggested

Key grouping

- **Required**: Data → Tokenization → BERT embeddings → Keras/TF modeling → Training → Recall-focused evaluation → TF Serving

- **Suggested**: Word2Vec/ELMo for extra features, HF Trainer to simplify loops, LIME/SHAP for transparency, HF Inference API for ops shortcuts.