Given:

- Labeled set: $\mathcal{L} = \{(x_i, y_i)\}_{i=1}^{2K}$ where K examples per class
- Unlabeled set: $\mathcal{U} = \{x_j\}_{j=1}^{M}$
- Test set: $\mathcal{T} = \{x_k\}_{k=1}^{N}$
- Constraints: $K \ll M, N$

Objective:

- Learn classifier $f : \mathcal{X} \to \{0, 1\}$ that accurately predicts labels for $\mathcal{T}$
- Binary classification: real news ($y = 0$) vs fake news ($y = 1$)

Key Challenges:

- Extreme data scarcity: $K \in \{3 \sim 16\}$ labeled examples per class
- Content-only constraint: No user interaction or propagation data available