

Methodology

Problem Definition and Key Notation

- Given a news article $A = \{T, V\}$ (T = text information, V = visual information)
- Denote $t, v \in \mathbb{R}^d$ as corresponding representations, $t = M_t(T, \theta_t)$, $v = M_v(V, \theta_v)$
- Let $s = M_s(t, v)$ denote the similarity between t and v , where $s \in [0, 1]$
- Goal: $M_p : (M_t, M_v, M_s) \xrightarrow{(\theta_t, \theta_v, \theta_p)} \hat{y} \in [0, 1]$, where θ_* are parameters to be learned
 - Determine whether A is fake news ($\hat{y} = 1$) or true one ($\hat{y} = 0$).
 - By investigating its textual, visual information, and their relationship.

Methodology

Multi-modal Feature Extraction - Text

- Extend Text-CNN by introducing an additional fully connected layer to automatically extract textual features for each news article.

