Text classification poste- Autoencoder reconstruc- Next word probability Other objectives rior probability tion P(C|D)P(X'|X) $P(X_t|X_{t-1})$ Task-Specific Representation Lanauge Model Text Classification Autoencoder Semantic Representation X: Input