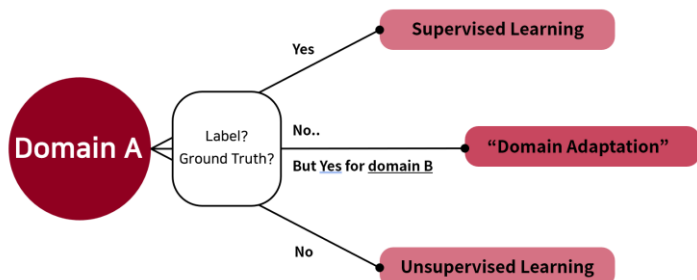




Introduction

Motivation

- 풍부한 영어 감성 데이터로 학습한 모델을 라벨이 부족한 한국어 데이터에 활용



Key Idea

- 영어 감성 데이터로 학습한 모델을 한국어 감성 분류에 domain adaptation

Dataset

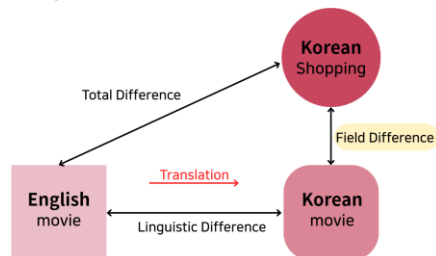
- 영어 텍스트 데이터(Label): IMDb, SST, Dynasent
- 한국어 텍스트 데이터(Unlabel): 네이버 쇼핑 리뷰 텍스트 데이터

Approach

- Translation-based: 라벨이 있는 English 데이터를 Korean으로 번역해 MMD loss로 학습
- Translation-free: DANN, FreeMatch를 이용해 English와 Korean 데이터 분포를 맞춰 한국어 성능 향상

Method

Objective



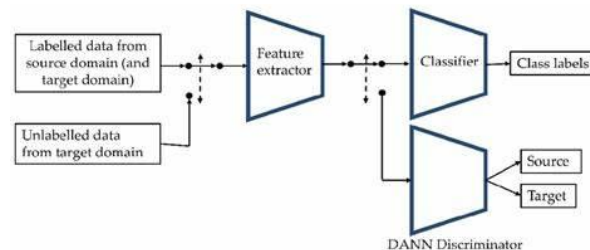
Model

- Translation-based: MMD(Maximum Mean Discrepancy)

$$\text{MMD}^2(P, Q) = \frac{1}{m^2} \sum_{i=1}^m \sum_{j=1}^m k(x_i, x_j) + \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n k(y_i, y_j) - \frac{2}{mn} \sum_{i=1}^m \sum_{j=1}^n k(x_i, y_j),$$

where $x_i \sim P, y_j \sim Q$

- Translation-free: DANN + Free Match



Fix-Match

Fixed Threshold
모델의 예측 신뢰도가 사전에 정의된 값(예: 0.95)을 넘을 때만 유사 레이블로 인정하고 학습에 사용

Free-Match

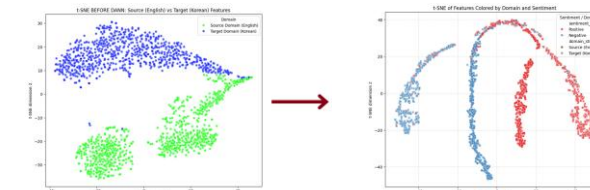
Self Adaptation Threshold
threshold값을 고정하지 않고, 모델의 학습 상태에 따라 매 순간 동적으로 조절

Result

Quantitative Results

	Accuracy	F1-score	AUC-ROC
Multi lingual direct	0.7948	0.8147	0.8949
Rich-resource case	0.9316	0.9327	0.9771
DANN	0.8648	0.8644	0.9354
DANN + Free-Match	0.9064	0.9064	0.9451
With MMD	0.9078	0.9078	0.9613

Qualitative Results



Conclusion

Summary of Findings

- DANN + FreeMatch 파이프라인**
 - DANN: 도메인 간 차이를 해소하여 안정적인 학습
 - FreeMatch: 타겟 도메인 데이터의 feature를 학습하여 성능 최적화
- Low resource 상황에서 Domain Adaptation으로 고성능 모델 구현 가능

Limitation & Future Work

- Task Extension: 다중 클래스 감성 분류
- 방법론 개선: Spectral Normalization