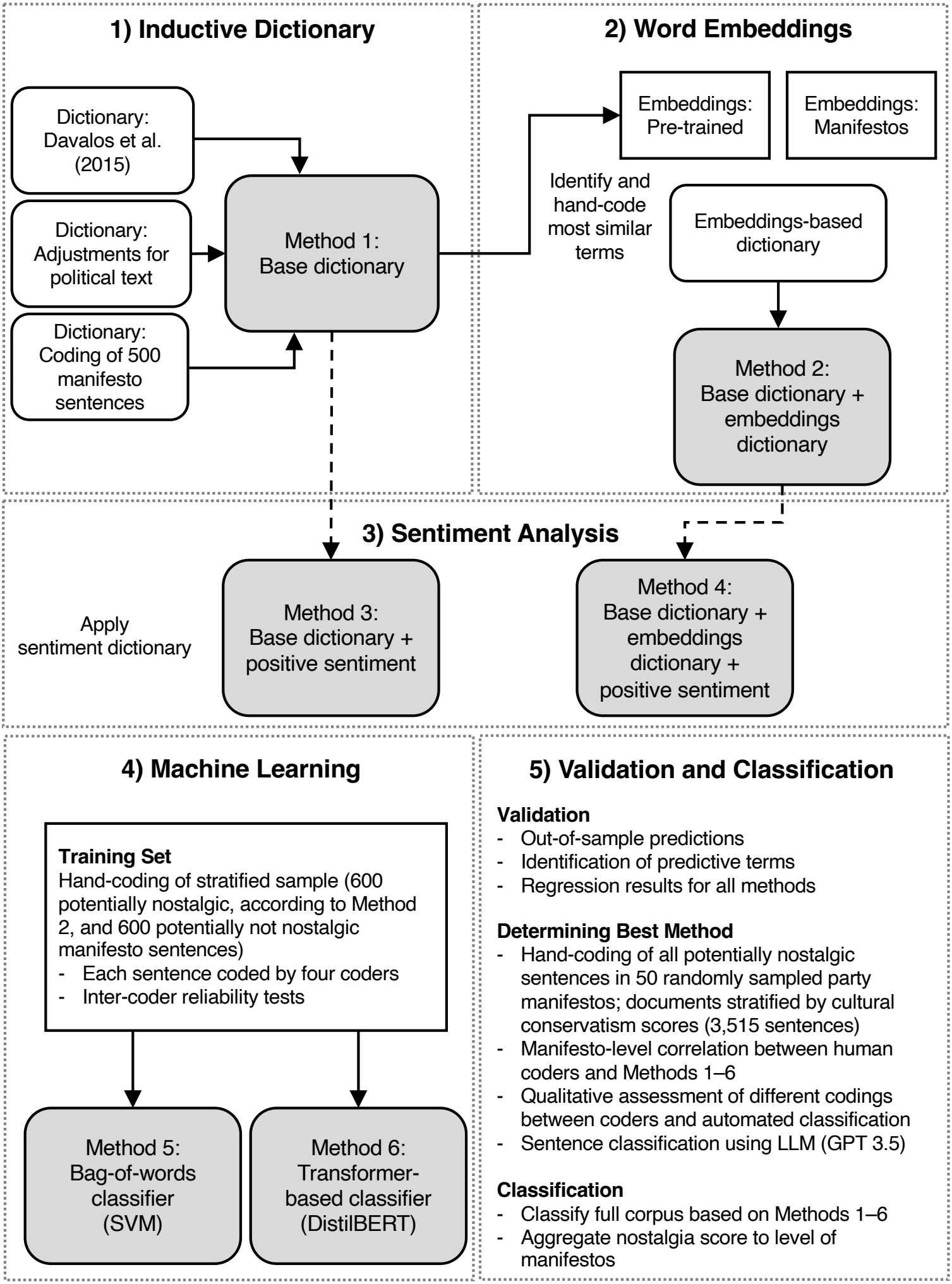
1) Inductive Dictionary
Dictionary: Davalos et al. (2015)
Dictionary: Adjustments for political text
Dictionary: Coding of 500 manifesto sentences
Method 1: Base dictionary
2) Word Embeddings
Embeddings: Pre-trained
Embeddings: Manifestos
Identify and hand-code most similar terms
Embeddings-based dictionary
Method 2: Base dictionary + embeddings dictionary
3) Sentiment Analysis
Apply sentiment dictionary
Method 3: Base dictionary + positive sentiment
Method 4: Base dictionary + embeddings dictionary + positive sentiment
4) Machine Learning
Training Set
Hand-coding of stratified sample (600 potentially nostalgic, according to Method 2, and 600 potentially not nostalgic manifesto sentences)
- Each sentence coded by four coders
- Inter-coder reliability tests
Method 5: Bag-of-words classifier (SVM)
Method 6: Transformer-based classifier (DistilBERT)
5) Validation and Classification
Validation
- Out-of-sample predictions
- Identification of predictive terms
- Regression results for all methods
Determining Best Method
- Hand-coding of all potentially nostalgic sentences in 50 randomly sampled party manifestos; documents stratified by cultural conservatism scores (3,515 sentences)
- Manifesto-level correlation between human coders and Methods 1–6
- Qualitative assessment of different codings between coders and automated classification
- Sentence classification using LLM (GPT 3.5)
Classification
- Classify full corpus based on Methods 1–6
- Aggregate nostalgia score to level of manifestos