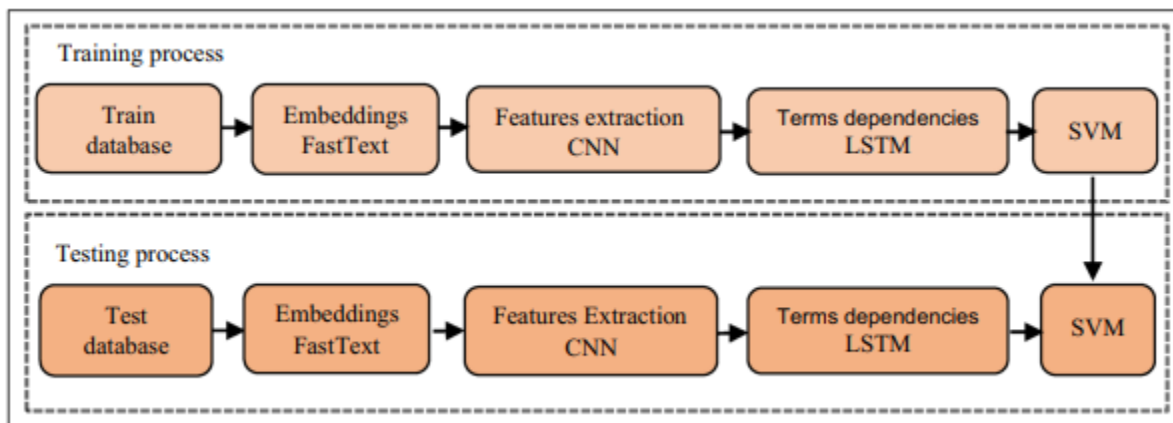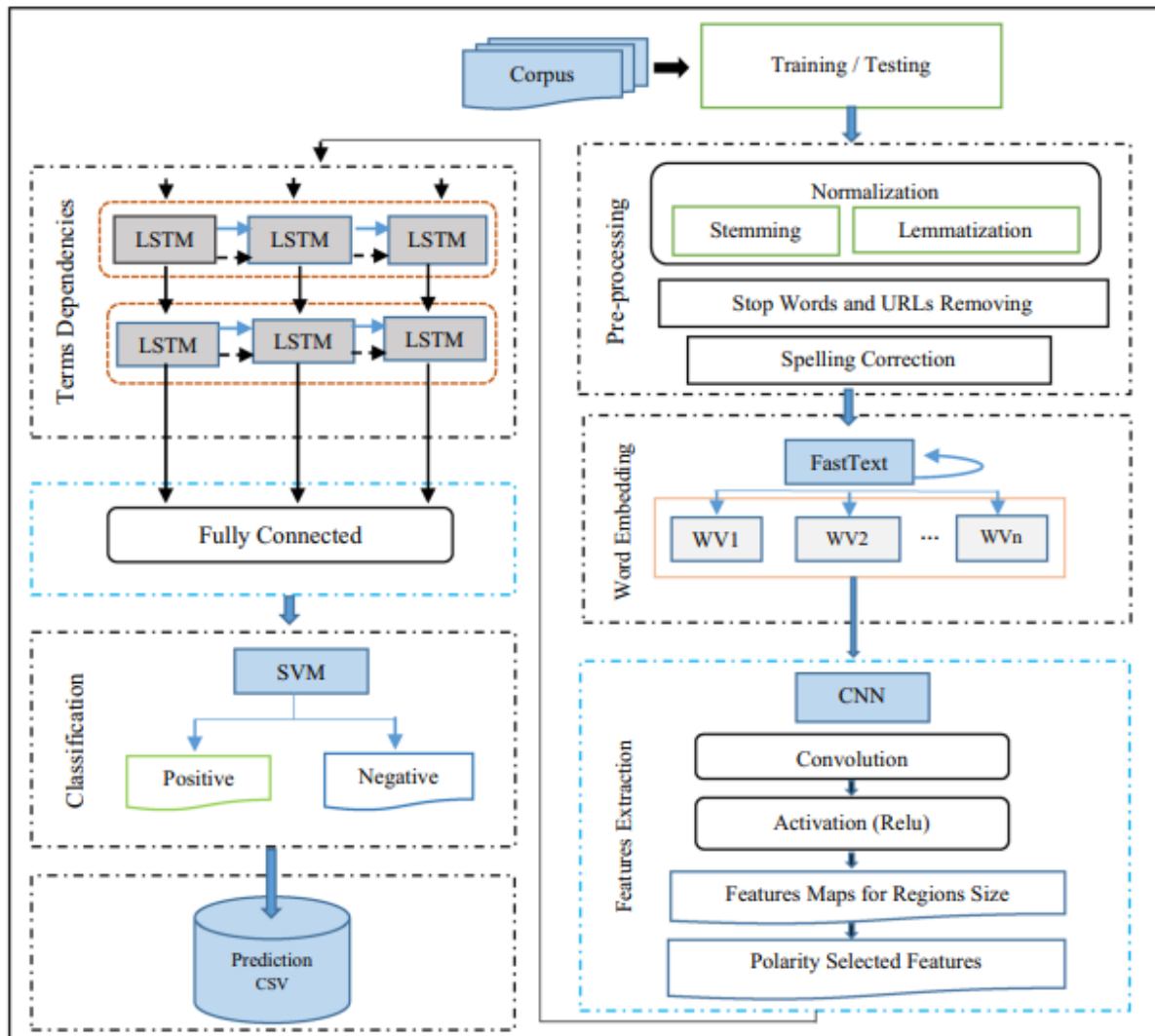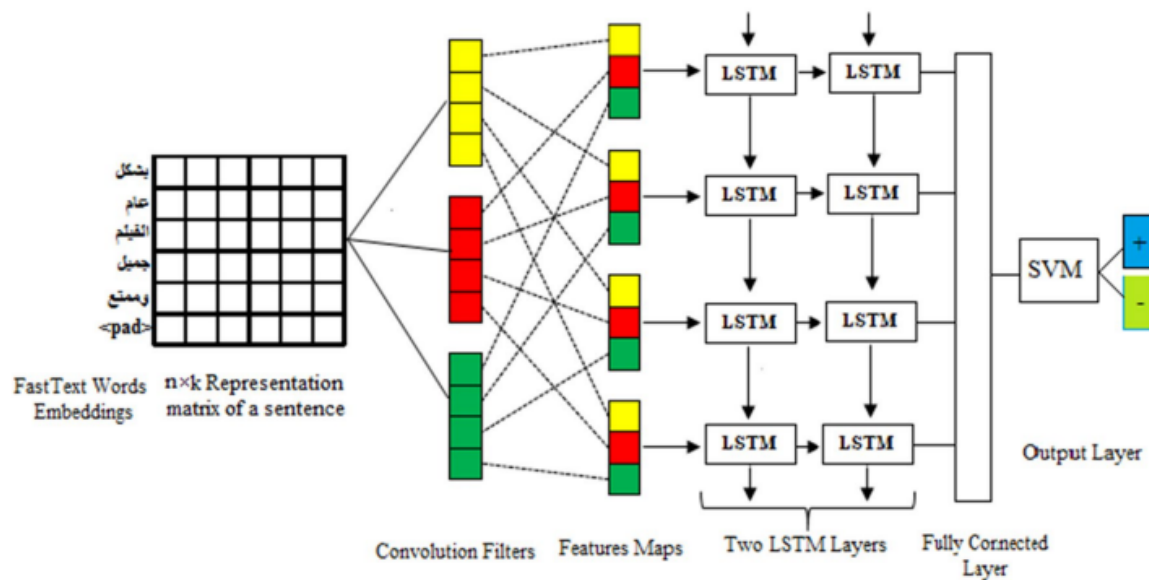# Deep learning CNN–LSTM framework for Arabic sentiment analysis using textual information shared in social networks

Pipeline:

Architecture:

Corpus → **Training / Testing**

Pre-processing
- Normalization
  - Stemming
  - Lemmatization
- Stop Words and URLs Removing
- Spelling Correction

Word Embedding
- FastText
  - WV1
  - WV2
  - ...
  - WVn

Features Extraction
- CNN
- Convolution
- Activation (Relu)
- Features Maps for Regions Size
- Polarity Selected Features

Terms Dependencies
- LSTM — LSTM — LSTM
- LSTM — LSTM — LSTM
- Fully Connected

Classification
- SVM
  - Positive
  - Negative

Prediction CSV

FastText Words Embeddings   n×k Representation matrix of a sentence   Convolution Filters   Features Maps   Two LSTM Layers   Fully Cornected Layer   Output Layer

Results:

Classifier	Precision (%)	Recall (%)	F1-score (%)	Accuracy (%)
SVM	89.10	92.14	92.44	90.75
NB	88.02	87.60	87.56	88.75
Softmax	87.81	86.63	87.39	87.43
KNN	86.72	86.94	86.73	86.83
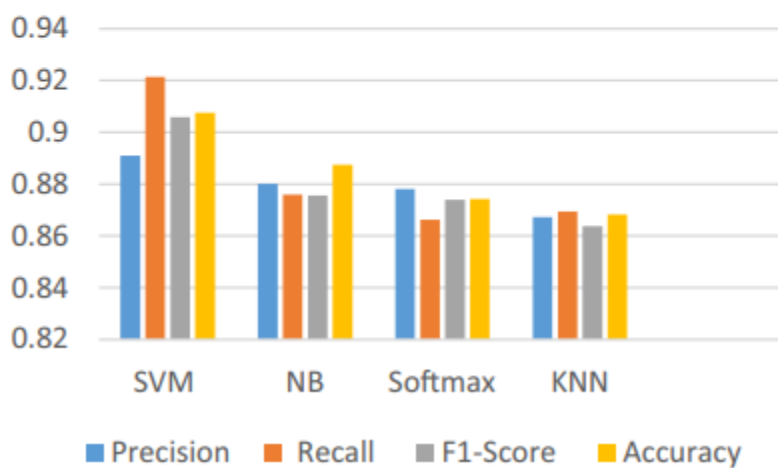
Table 6 Effects of the number of LSTM layers

N. of LSTM layers	Precision (%)	Recall (%)	Accuracy (%)
1	86.10	89.45	87.98
2	89.10	92.14	90.75

Table 7 Classification accuracy using different embedding models

Embedding model	Classification accuracy (%)
FastText(CBOW)-CNN–LSTM	88.90
FastText(Skip-gram)-CNN–LSTM	90.75
Word2Vec-CNN–LSTM	86.84
AraVec-CNN–LSTM	81.35