

GuessMyMood

Text-based Emotion Classification using
two-pass Neural Network Architecture

Navi Tansaraviput, Laphonchai Jirachuphun, Varot Premtoon



Overview

Task: Classify each sentence in the dataset into one of the seven emotions:

1. NEUTRAL
2. HAPPY
3. SURPRISED
4. SAD
5. ANGRY
6. FEARFUL
7. DISGUSTED



Dataset

Corpus: EBBA CECILIA OVESDOTTER ALM

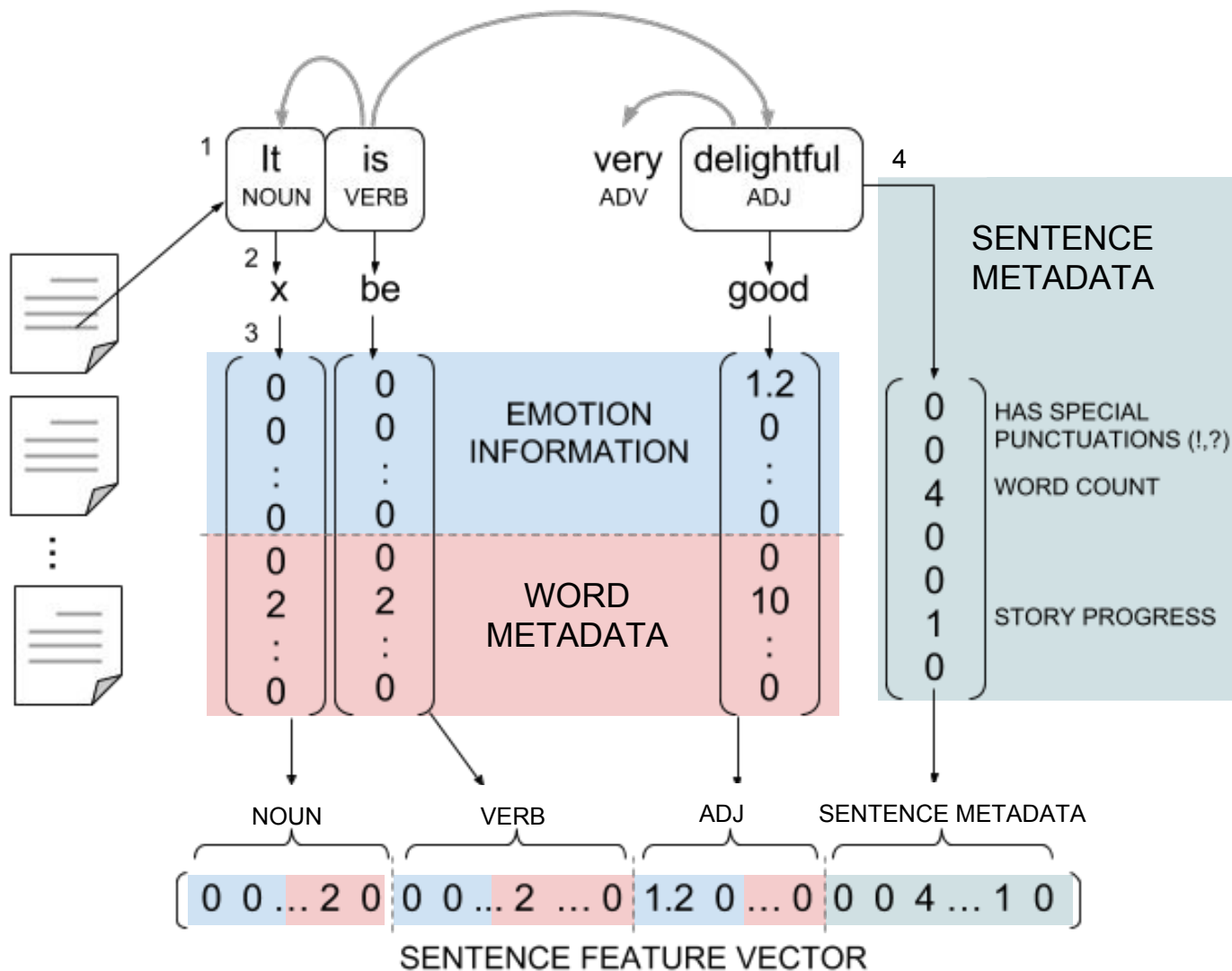
- Annotated 176 children stories
- A total of 15,302 sentences
- Two annotators

1st annotator is used as gold standard

Percentage of each emotion in corpus

| Neutral | Happy | Surprised | Sad | Angry | Fearful | Disgusted |
|---------|-------|-----------|------|-------|---------|-----------|
| 66.3% | 10.5% | 5.4% | 5.4% | 4.8% | 4.6% | 3.0% |

Feature extraction



Feature extraction

LEMMATIZING "DELIGHTFUL"



EMOTION
INFORMATION

$$\begin{matrix}
 \text{POSITIVE} & \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \\
 \text{HAPPY} & \\
 \text{TRUST} &
 \end{matrix}
 \times
 \begin{matrix}
 0.8 \\
 \text{SIMILARITY} \\
 \text{(delightful, good)}
 \end{matrix}
 =
 \begin{bmatrix} 0.8 \\ 0 \\ 0 \\ 0.8 \\ 0 \\ 0 \\ 0.8 \\ 0 \end{bmatrix}$$

DEPENDENCY TREE FOR "DELIGHTFUL"



DEGREE WORD
INTENSITY
DICTIONARY

INTENSITY

1.5
VERY

$$\begin{matrix}
 \begin{bmatrix} 0.8 \\ 0 \\ 0 \\ 0.8 \\ 0 \\ 0 \\ 0.8 \\ 0 \end{bmatrix} \\
 \text{DEGREE WORD} \\
 \text{INTENSITY} \\
 \text{DICTIONARY}
 \end{matrix}
 \times
 \begin{matrix}
 1.2 \\
 0 \\
 0 \\
 1.2 \\
 0 \\
 0 \\
 1.2 \\
 0
 \end{matrix}
 =
 \begin{bmatrix} 1.2 \\ 0 \\ 0 \\ 1.2 \\ 0 \\ 0 \\ 1.2 \\ 0 \end{bmatrix}$$

FEATURE VECTOR
FOR "DELIGHTFUL"

$$\begin{bmatrix} 1.2 \\ 0 \\ : \\ 0 \\ 0 \\ 10 \\ : \\ 0 \end{bmatrix}$$

WORD METADATA

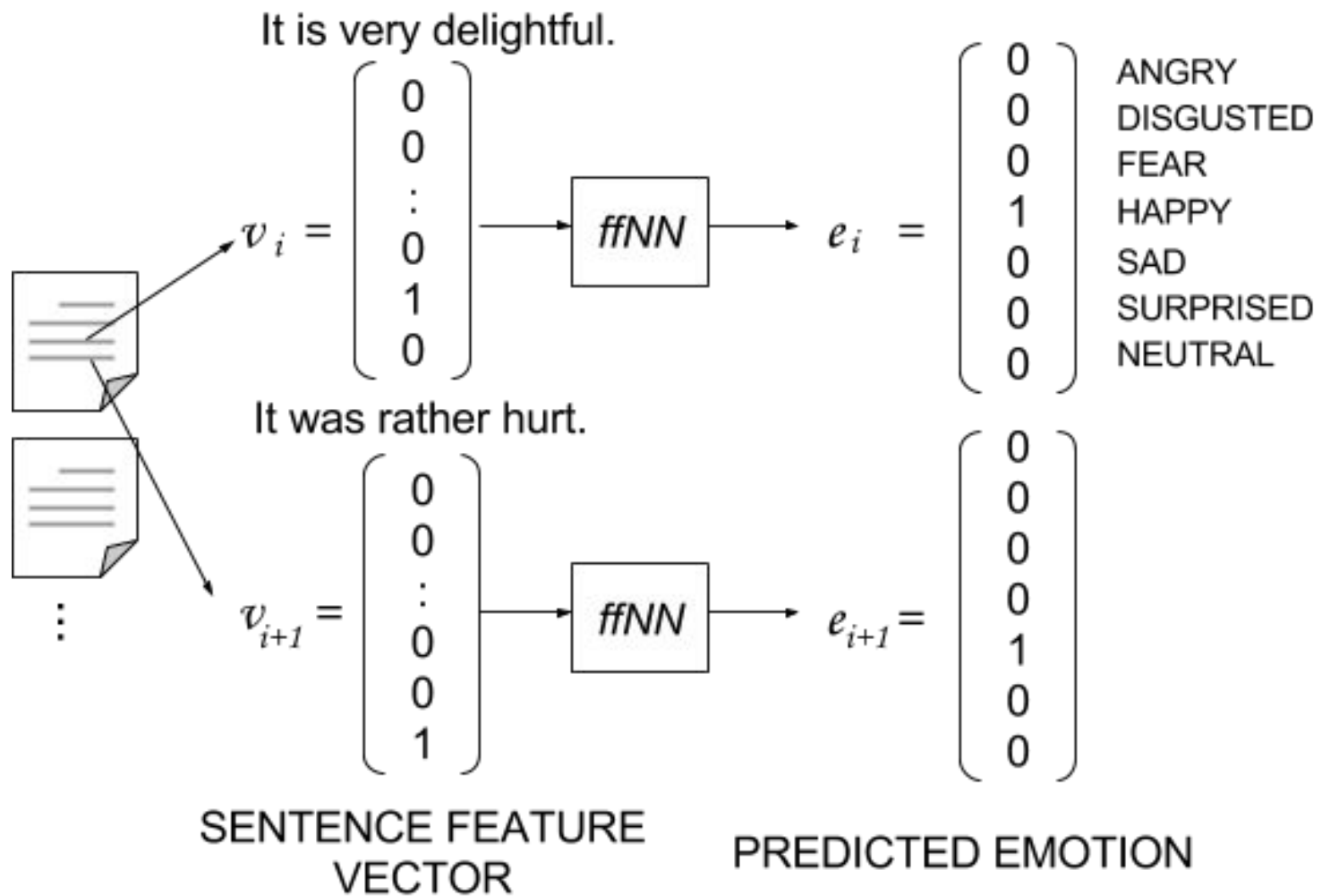
delightful
ADJ

IS UPPERCASE
WORD LENGTH

PART OF SPEECH
(ADJ)

$$\begin{bmatrix} 0 \\ 10 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

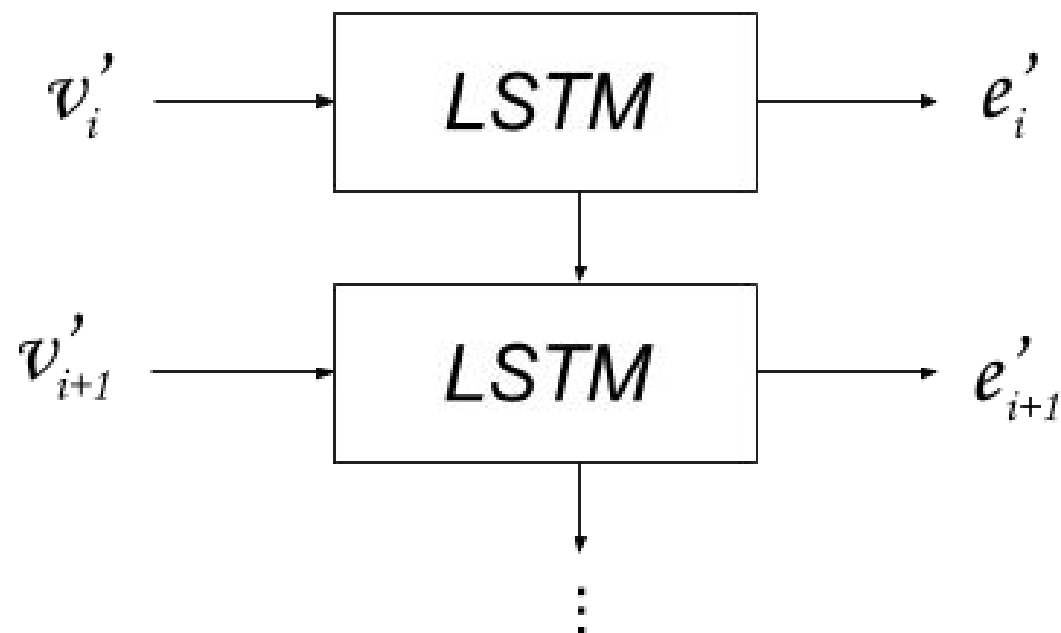
Model



Model

AUGMENTED FEATURE VECTOR

$$\mathbf{v}'_i = \left[\mathbf{e}_{i-2} \ \mathbf{e}_{i-1} \ \mathbf{e}_i \ \mathbf{e}_{i+1} \ \mathbf{e}_{i+2} \mid \mathbf{v}_i \right]$$



Results

| Model | Precision | Recall | F1 |
|---------------------|--------------|--------------|--------------|
| 2nd human annotator | 58.2% | 40.3% | 45.6% |
| Baseline | | | |
| random guess | 14.3% | 13.4% | 9.8% |
| all neutral | 9.5% | 14.3% | 11.4% |
| Model | | | |
| Decision Tree | 17.9% | 17.9% | 17.8% |
| SVM | 16% | 16.2% | 13.2% |
| 1-pass ffNN | 21.6% | 25.8% | 20.4% |
| 2-pass ffNN + LSTM | 16.6% | 14.3% | 11.5% |
| 2-pass ffNN + ffNN | 20.2% | 23% | 18% |