

CSE691 Assignment 4 Report

Lichen Liang

1. The long text has the following output:

```
First Layer Embeddings:  
Closest Words to Test Word (she): ['she', 'know--no', 'crown', 'parchment', 'prove', 'beasts']  
Closest Words to Test Word (queen): ['queen', 'n', 'corner', 'thick', '_please_', 'sharply']  
Second Layer Embeddings:  
Closest Words to Test Word (she): ['she', 'and', 'alice', 'to', 'was', 'it']  
Closest Words to Test Word (queen): ['queen', 'went', 'me', 'here', 'quite', 'about']
```

The short text has the following output:

```
First Layer Embeddings:  
Closest Words to Test Word (she): ['she', 'the', 'he', 'queen', 'king', 'is']  
Closest Words to Test Word (queen): ['queen', 'he', 'she', 'the', 'is', 'king']  
Second Layer Embeddings:  
Closest Words to Test Word (she): ['she', 'he', 'king', 'royal', 'queen', 'the']  
Closest Words to Test Word (queen): ['queen', 'king', 'he', 'she', 'the', 'royal']
```

2. At lowest loss, the weights of the first layer are shown in figure 1.

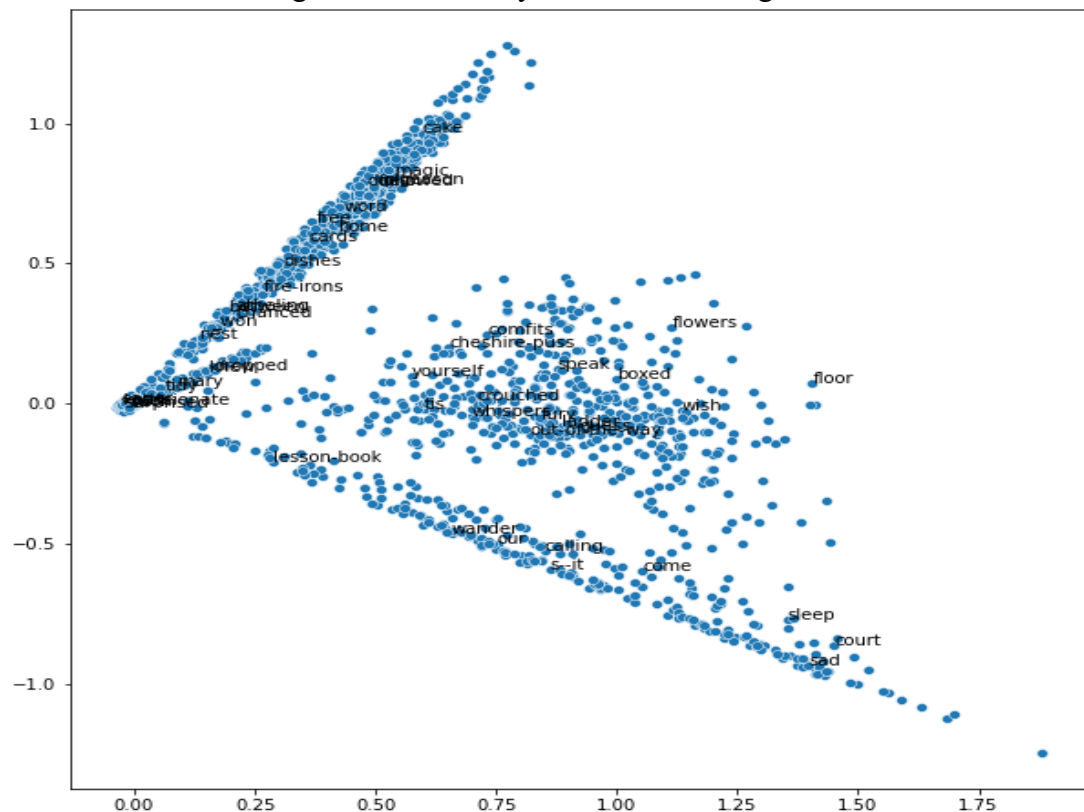


Figure 1.

At lowest loss, the weights of the second layer are shown in figure 2.

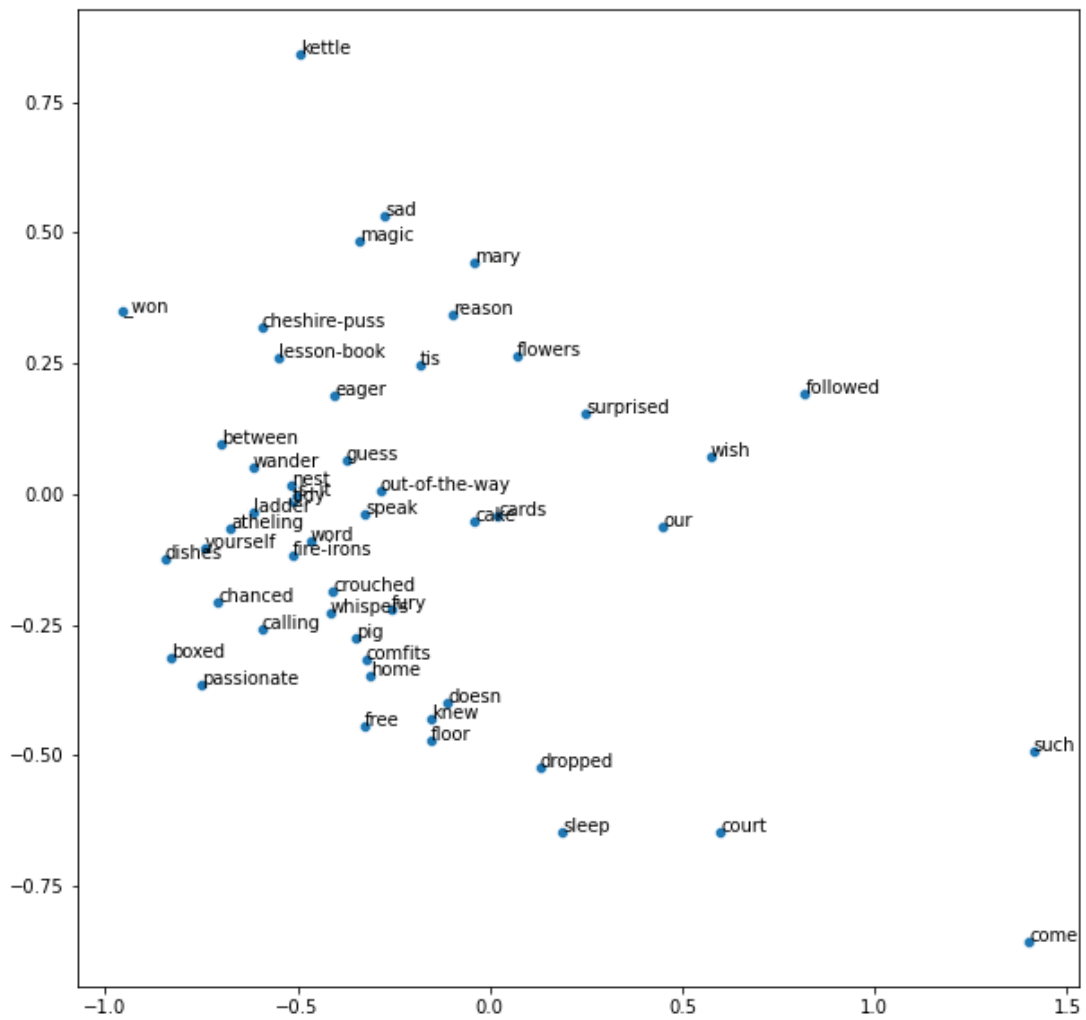


Figure 2.

As we can see, the number of words after the second layer is much less than the first layer. Layer two is better than layer one as it filters out more words and narrows down on the result.