Sentiment Analysis with Deep Learning

Andisa Dewi
andisa.dewi@knime.com
Sentiment Analysis – An Example

Samsung

Samsung Galaxy S7 Edge G935A 32GB Unlocked - Gold Platinum

4.9 stars | 125 customer reviews | 606 answered questions

🌟🌟🌟🌟🌟 Beautiful phone from a wonderful seller!
By ay on May 29, 2017
Color: Gold | Verified Purchase

This practically new beautiful phone well exceeded my expectations!

🌟🌟🌟🌟 One Star
By on August 3, 2016
Color: Black Onyx | Verified Purchase

Very bad experience
Sentiment Analysis

Task: Determine the expressed opinion in a document/text, e.g. positive, negative

Sentiment Analysis = Opinion Mining = Emotion AI

Lexicon Based

Machine Learning

Deep Learning
Philosophy

Reading/Parsing Data

Preprocessing

Transformation

Classification

... perhaps your name is Rumpelstiltskin?

© 2020 KNIME AG. All rights reserved.
Part 1: Reading and Parsing Data

Read/Parse textual data

<table>
<thead>
<tr>
<th>sentiment</th>
<th>text</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Although the production and Jerry Jameson's direction</td>
</tr>
<tr>
<td>0</td>
<td>Capt. Gallagher (Lemmon) and flight attendant Eve C</td>
</tr>
<tr>
<td>0</td>
<td>Towards the end of the movie, I felt it was too techni</td>
</tr>
<tr>
<td>0</td>
<td>This is the kind of movie that my enemies content I w</td>
</tr>
<tr>
<td>0</td>
<td>I saw 'Descent' last night at the Stockholm Film Festi</td>
</tr>
<tr>
<td>0</td>
<td>Some films that you pick up for a pound turn out to b</td>
</tr>
<tr>
<td>0</td>
<td>This is one of the dumbest films, I've ever seen. It ri</td>
</tr>
<tr>
<td>1</td>
<td>Bromwell High is a cartoon comedy. It ran at the same</td>
</tr>
<tr>
<td>1</td>
<td>Homelessness (or Houselessness as George Carlin sai</td>
</tr>
<tr>
<td>1</td>
<td>Brilliant over-acting by Lesley Ann Warren. Best dram</td>
</tr>
<tr>
<td>1</td>
<td>This is easily the most underrated film inn the Brooks</td>
</tr>
<tr>
<td>1</td>
<td>This is not the typical Mel Brooks film. It was much le</td>
</tr>
<tr>
<td>1</td>
<td>This isn't the comedic Robin Williams, nor is it the qu</td>
</tr>
<tr>
<td>1</td>
<td>Yes its an art... to successfully make a slow paced th</td>
</tr>
<tr>
<td>1</td>
<td>In this critically acclaimed psychological thriller ba</td>
</tr>
<tr>
<td>1</td>
<td>THE NIGHT LISTENER (2006) **1/2 Robin Williams, Tor</td>
</tr>
<tr>
<td>1</td>
<td>You know, Robin Williams, God bless him, is constant</td>
</tr>
<tr>
<td>1</td>
<td>When I first read Armistead Maupins story I was take</td>
</tr>
<tr>
<td>1</td>
<td>I liked the film. Some of the action scenes were very</td>
</tr>
</tbody>
</table>
Part 2: Preprocessing

Example:
- This movie is horrible. The acting is a waste basket. Though the scenery is great.
- Even though this movie came out a year before I was born, it’s my favorite movie.
- This is definitely one of my favorite comedies.

- movie horrible acting waste basket scenery
- movie born favorite movie
- definitely favorite comedies
Part 3: Transformation

Transform documents to numbers

<table>
<thead>
<tr>
<th>Row ID</th>
<th>D movi</th>
<th>D absolut</th>
<th>D act</th>
<th>D worst</th>
<th>D amaz</th>
<th>D aw</th>
<th>D direct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Row0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Row1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Row2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Row3</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Row4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Row5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Row6</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Row7</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Row8</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Row9</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Row10</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Row11</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Row12</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Row13</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Row14</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Transformation for Deep Learning

Expected input of a network:

- Numerical representation of each document encoding the words and their order
- Equivalent input shape of each document
  - Truncate too long documents
  - Zero pad too short documents
Transformation Example

- movie horrible acting waste basket scenery
- movie born favorite movie
- definitely favorite comedies

<table>
<thead>
<tr>
<th>Term</th>
<th>Int</th>
</tr>
</thead>
<tbody>
<tr>
<td>favorite</td>
<td>1</td>
</tr>
<tr>
<td>movie</td>
<td>2</td>
</tr>
<tr>
<td>acting</td>
<td>3</td>
</tr>
<tr>
<td>basket</td>
<td>4</td>
</tr>
<tr>
<td>born</td>
<td>5</td>
</tr>
<tr>
<td>comedies</td>
<td>6</td>
</tr>
<tr>
<td>definitely</td>
<td>7</td>
</tr>
<tr>
<td>horrible</td>
<td>8</td>
</tr>
<tr>
<td>scenery</td>
<td>9</td>
</tr>
<tr>
<td>waste</td>
<td>10</td>
</tr>
</tbody>
</table>

- [2 8 3 10 4 9]
- [2 5 1 2]
- [7 1 6]

- [2 8 3 10 4]
- [2 5 1 2]
- [7 1 6]

- [2 8 3 10 4]
- [2 5 1 2 0]
- [7 1 6 0 0]
Transformation for Deep Learning

[Diagram showing data flow and operations such as Table Reader, Strings To Document, Preprocessing, Dictionary Filter, Truncate Dictionary Replacer, Zero Pad, etc.]

[Table showing data with columns for Row ID, sentiment, AggregatedValues]
Part 4: Classification

Define Network
- Keras Input Layer
  - Shape: max number of words per documents
- Keras Embedding Layer
  - Input: # word in dictionary + 1
  - 128 units
- Keras LSTM Layer
- Keras Dense Layer
  - Activation function: sigmoid

Reading and Preprocessing
- Table Reader
- Strings To Document
- Preprocessing
- Dictionary Filter
- Truncate Dictionary Replacer
- Zero Pad
- Partitioning
- Table Reader
- Reduce Dictionary
  - vocab: 20000

Training and Predicting
- Keras Network Learner
  - Train the network
- Keras Network Executor
  - Predict test data

Evaluation
- Extract prediction Scorer (JavaScript)
Neural Network
Neural Network
Neural Network
Neural Network
Deep Learning

Artificial neural network

Deep neural network

Input  Hidden  Output

Input  Hidden  Hidden  Output

n hidden layers
(convolutional, embedding, dense, recurrent, ...)

© 2020 KNIME AG. All rights reserved.
Part 4: Classification
Our Network Topology
Input Layer

- Input layer passes input data to the first hidden layer
- The dimension is the document size, i.e. the number of words in each document
**Embedding Layer**

- Embedding layer embeds each word into a dense, high dimensional vector

- Example:

- Similar words will be embedded near each other in the vector space
RNN (Recurrent Neural Network)

- RNN is a deep learning model
- It has the ability to memorize previous inputs
- Suitable for analysing sequential data, e.g. text, time series
- However, unreliable in handling long-term memory
LSTM (Long Short Term Memory) Layer

- LSTM is a variant of RNN
- Handles long-term memory
- Uses gates to control memorizing process

Example:

\[
\begin{align*}
\{[0.3, 0.2, \ldots, 0.6]\} & \rightarrow \{0.3, 0.4, 0.7, \ldots\} \\
\{[0.4, 0.6, \ldots, 0.5]\} & \rightarrow \{0.3, 0.4, 0.7, \ldots\} \\
\{[0.1, 0.8, \ldots, 0.3]\} & \rightarrow \{0.3, 0.4, 0.7, \ldots\}
\end{align*}
\]
Dense Layer

- Dense layer connects each unit of the input with each output unit of this layer.
- Defines the activation function for the final output.

Example:

\[
p = \begin{pmatrix} 0.3 \\ 0.4 \\ 0.7 \\ \cdots \end{pmatrix} \rightarrow \begin{pmatrix} 0.8 \end{pmatrix}
\]
KNIME Deep Learning

- **Keras Integration**
Part 4: Classification
Training the Network

Keras Network Learner

Train the network

Dialog - 0:189 - Keras Network Learner (train for 3 epochs)

General Settings
- Back end: Keras (TensorFlow)
- Epochs: 3
- Training batch size: 32
- Validation batch size: 32

Optimizer Settings
- Optimizer: Adam
- Learning rate: 0.001
- Beta 1: 0.9
- Beta 2: 0.999
- Epsilon: 1.0E-8
- Learning rate decay: 0.0
- Clip norm: 1.0
- Clip value: 1.0
Training the Network: Learning Monitor
Dataset

• Subset of the IMDb (Internet Movie Database) Large Movie Review Dataset v1.0 with 50000 documents*.
  – 25000 documents from the positive group
  – 25000 documents from the negative group

• Goal: Assign the correct sentiment label to each document.

(*) For details about the data set see http://ai.stanford.edu/~amaas/data/sentiment/
Data citation: Andrew L. Maas, Raymond E. Daly, Peter T. Pham, Dan Huang, Andrew Y. Ng, and Christopher Potts. (2011). Learning Word Vectors for Sentiment Analysis. The 49th Annual Meeting of the Association for Computational Linguistics (ACL 2011)
Setting up the Keras Integration

• Install the KNIME Keras Integration extension
  – Go to File > Install KNIME Extensions…
  – Enter keras into the search box
  – The extension is listed under KNIME Labs Extensions

• Setup Python for KNIME Deep Learning using Anaconda
  – Python environment with Keras and TensorFlow

Please follow the installation details in the KNIME Keras Integration Installation guide:

https://docs.knime.com/latest/deep_learning_installation_guide/index.html
References

Word Embedding:  
https://www.knime.com/blog/word-embedding-word2vec-explained

RNN/LSTM:  
https://www.knime.com/blog/text-generation-with-lstm

KNIME Deep Learning - Keras Integration  
https://www.knime.com/deeplearning/keras

KNIME Deep Learning - TensorFlow Integration  
https://www.knime.com/deeplearning/tensorflow
Demo
The End – thank you for joining this webinar.