

TensorFlow

Applications for natural language processing

COT-6930 Natural Language Processing
Spring 2019
Christian Garbin

What we will cover today

What TensorFlow is

A brief introduction to TensorFlow

Motivation to use TensorFlow

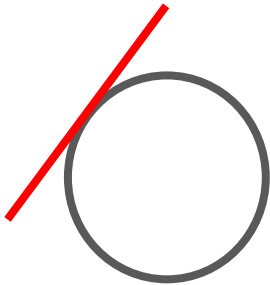
Applications for natural language processing

TensorFlow and natural language processing

TensorFlow

- How it works
 - The graph model behind TensorFlow
- What makes it unique among other deep learning frameworks
- The TensorFlow "family": TensorFlow, TensorFlow.js, TensorFlow lite
 - What is the application of each one
- The TensorFlow environment
 - Model visualization, optimization and debugging: TensorBoard
 - Reusable, pre-trained models: TensorFlow Hub
 - Example: universal sentence encoder:
<https://tfhub.dev/google/universal-sentence-encoder/2>
- TensorFlow for natural language processing

Going off on a (short) tangent



What is a tensor?

A generic name for n-dimensional data structures

Scalar
0D tensor

Rank = 0

'a'

Vector
1D tensor

Rank = 1

| |
|-----|
| 'a' |
| 'b' |
| 'c' |

Matrix
2D tensor

Rank = 2

| | |
|-----|-----|
| 'a' | 'x' |
| 'b' | 'y' |
| 'c' | 'z' |

Tensor
(n)D tensor



Rank = n

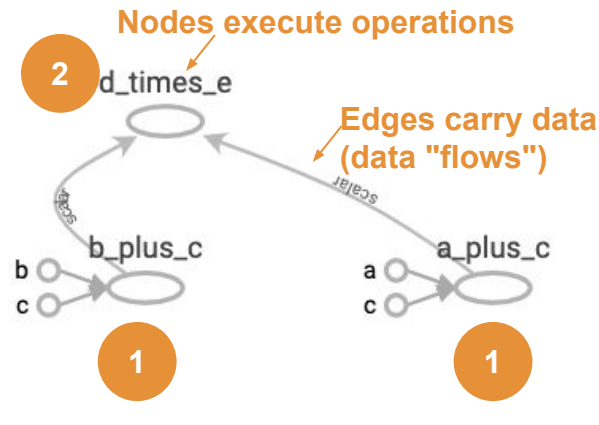
| | | | |
|-----|-----|-----|-----|
| | | 'a' | 'x' |
| | 'a' | 'b' | 'y' |
| 'a' | 'b' | 'c' | 'z' |
| 'b' | 'c' | | |
| 'c' | | | |

What is TensorFlow?

It is a generic, high-performance computation engine

It represents computations as a graphs

| | |
|---|--|
| a = 2 | a = tf.constant(2.0, name='a') |
| b = 2 | b = tf.constant(2.0, name='b') |
| c = 1 | c = tf.constant(1.0, name='c') |
|  | |
| d = b + c | d = tf.add(b, c, name='b_plus_c') |
| e = a + c | e = tf.add(a, c, name='a_plus_c') |
|  | |
| result = d * e | result = tf.multiply(d, e, name='d_times_e') |



Why "generic computation engine" matters

We can define and "package" higher-level operations

They can be efficiently distributed for parallel execution

They can be reused in different environments

Let's look at another example...

High-level ops "packaged"

The building block of neural networks

- A neuron: input * weight = output (*)
- Optimized with SGD (stochastic gradient descent)

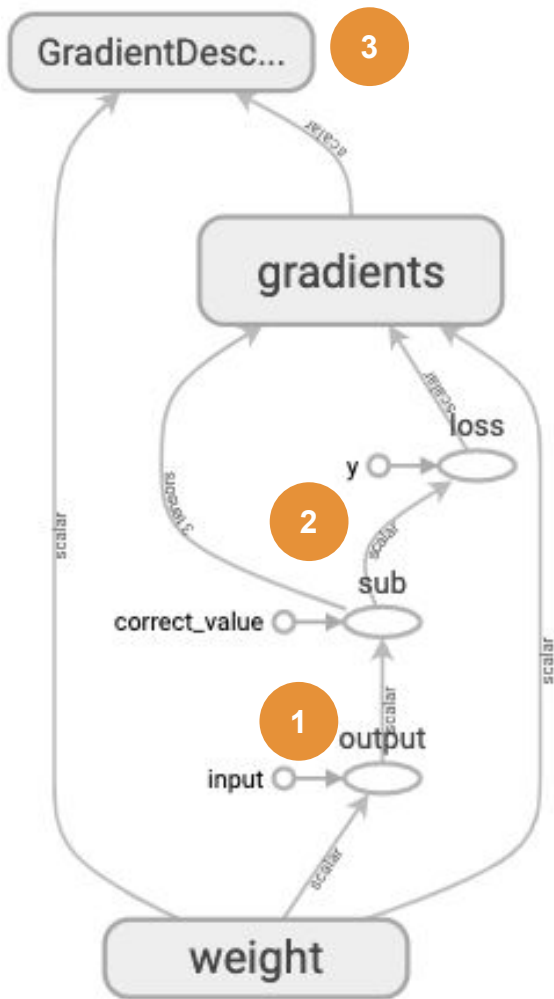


```
x = tf.constant(1.0, name='input')  
w = tf.Variable(0.8, name='weight')  
y = tf.multiply(w, x, name='output')
```

```
y_ = tf.constant(-1.0, name='correct_value')  
loss = tf.pow(y - y_, 2, name='loss')
```

```
train_step =  
    tf.train.GradientDescentOptimizer(0.025).minimize(loss)
```

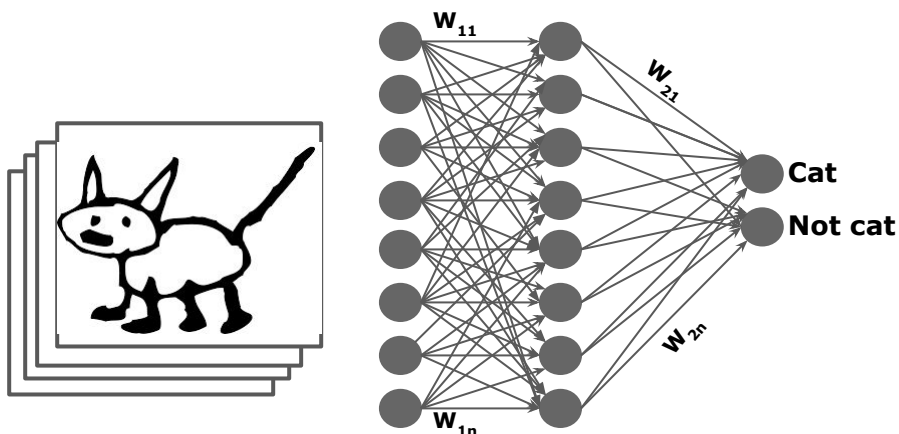
Higher-level operation



(*) Technically, this is missing a non-linearity after the output to be a neuron. It's simplified for illustration.

Why "high-performance" matters

Neural networks have millions of connections (parameters) and are trained on millions of samples



This results in billions of floating point matrix operations

$$(W^{l+1})^T \delta^{l+1} = \begin{bmatrix} w_{11}^{l+1} & \dots & w_{m1}^{l+1} & \dots & w_{M1}^{l+1} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ w_{1k}^{l+1} & \dots & w_{mk}^{l+1} & \dots & w_{Mk}^{l+1} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ w_{1K}^{l+1} & \dots & w_{mK}^{l+1} & \dots & w_{MK}^{l+1} \end{bmatrix} \begin{bmatrix} \delta_1^{l+1} \\ \vdots \\ \delta_m^{l+1} \\ \vdots \\ \delta_M^{l+1} \end{bmatrix} = \begin{bmatrix} \sum_{m=1}^M \delta_m^{l+1} w_{m1}^{l+1} \\ \vdots \\ \sum_{m=1}^M \delta_m^{l+1} w_{mk}^{l+1} \\ \vdots \\ \sum_{m=1}^M \delta_m^{l+1} w_{mK}^{l+1} \end{bmatrix}$$

TensorFlow is more than a library

TensorFlow Hub: pretrained models

TensorBoard: debugging and visualization tools

Deployment models: TensorFlow, TensorFlow.js, TensorFlow Lite

Pretrained models TensorFlow Hub

A collection of ready-to-use
models for several domains

Pick one and start using it, or
fine tune for your application

TensorFlow Hub

Text

Embedding

Image

Classification

Feature Vector

Generator

Other

Video

Classification

Publishers

Google

DeepMind

Text embedding



universal-sentence-encoder By Google

text-embedding DAN English

Encoder of greater-than-word length text trained on a variety of data.



universal-sentence-encoder-large By Google

text-embedding Transformer English

Encoder of greater-than-word length text trained on a variety of data.



elmo By Google

text-embedding 1 Billion Word Benchmark ELMo English

Embeddings from a language model trained on the 1 Billion Word Benchmark.

[View more text embeddings](#)

Image feature vectors



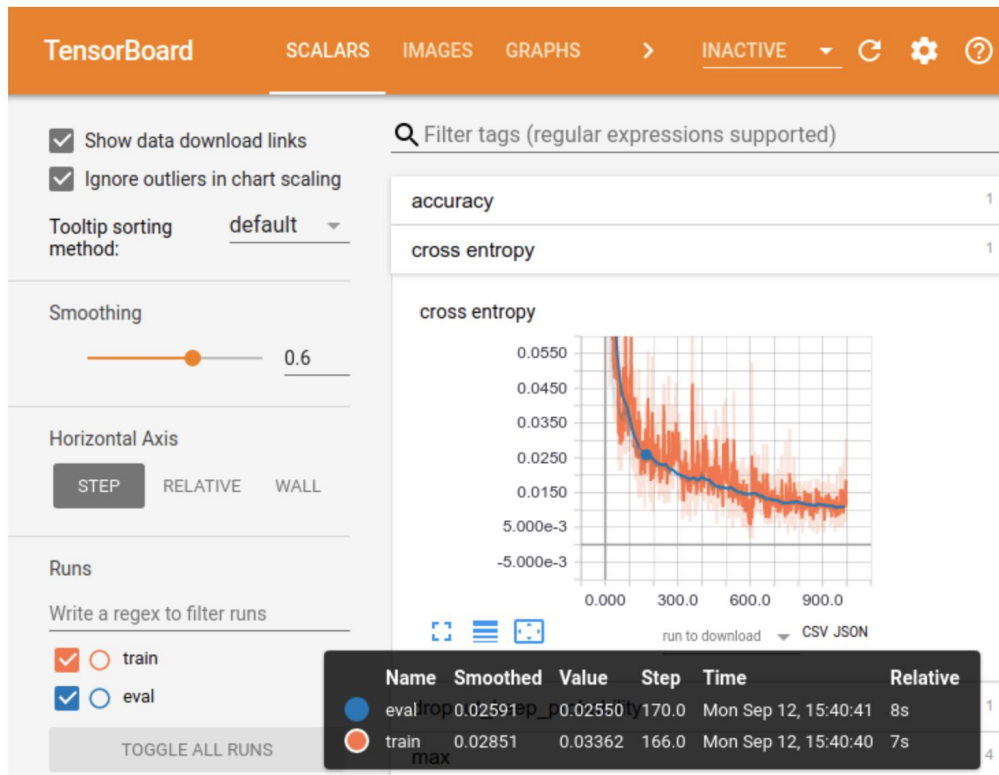
imagenet/inception_v3/feature_vector By Google

image-feature-vector ImageNet (ILSVRC-2012-CLS) Inception V3

Debugging and Visualization TensorBoard

Visualize and debug machine
learning

Perform "what if" analysis
without full retraining



Deployment models

TensorFlow.js, TensorFlow Lite

TensorFlow.js: run TensorFlow training and models inside a browser (or Node.js)

TensorFlow Lite: models optimized to run in low-powered devices (smartphones, smart appliances, etc.)

Putting it all together...

TensorFlow is...

- ...a generic computation engine that supports highly-parallelized execution on specialized hardware (GPUs, TPUs)
- ...adapted to run on diverse environments (TensorFlow, TensorFlow.js, TensorFlow Lite)
- ...backed by a collection of pretrained models (TensorFlow Hub) and a visualization/debugging tool (Tensorboard)

Back from the tangent...

TensorFlow and Natural Language Processing

TensorFlow and NLP

Utilities

Utility functions for natural language processing

Word and sentence embedding

Pretrained word and sentence embeddings for natural language processing

Visualizations

Tensorboard visualizations specific for NLP

TensorFlow utility functions

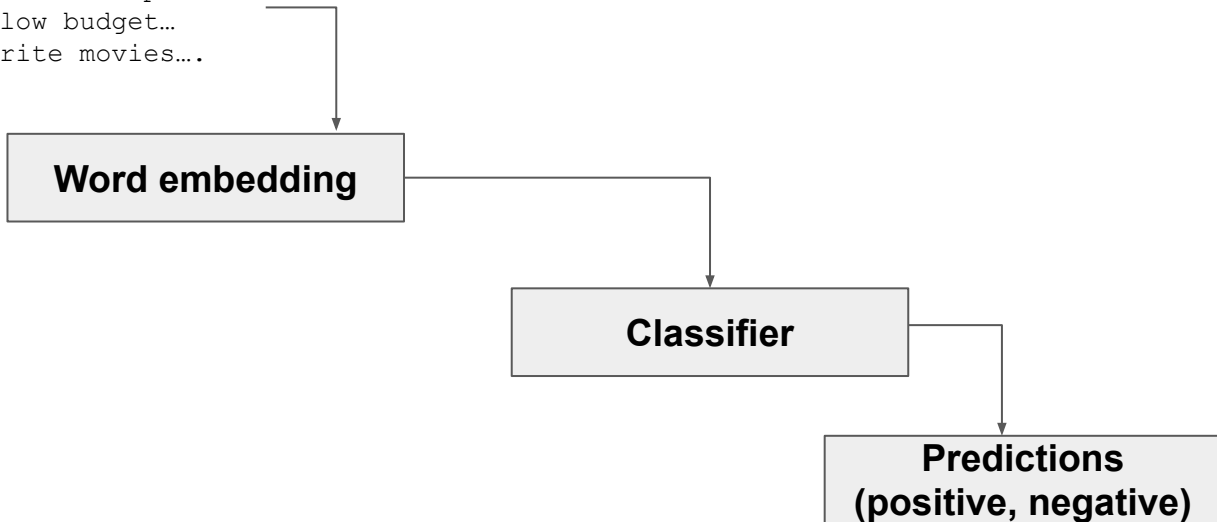
All concepts we learned in this class are available in TensorFlow

- From the simple ones
 - [tokenizer](#), [vocabulary creation](#), [n-grams](#), [tf-idf](#)
- To more complex ones:
 - [word \(text\) embedding](#), [word2vec representations](#)

TensorFlow NLP example

Movie sentiment classifier - the high level view

With the mixed reviews this got I wasn't expecting...
This film has a lot of raw potential. The script is...
Cage (1989) was another one of those low budget...
Home Alone 3 is one of my least favourite movies...



TensorFlow NLP example

Movie sentiment classifier - TensorFlow code

With the mixed reviews this got I wasn't expecting...
This film has a lot of raw potential. The script is...
Cage (1989) was another one of those low budget...
Home Alone 3 is one of my least favourite movies...

Word embedding

```
# pretrained (on Google News) word embedding (also tokenizes)
```

```
embedded_text_feature_column = hub.text_embedding_column(  
    key="sentence",  
    module_spec="https://tfhub.dev/google/nnlm-en-dim128/1")
```

```
# Neural network classifier, with two layers
```

```
estimator = tf.estimator.DNNClassifier(  
    hidden_units=[500, 100],  
    feature_columns=[embedded_text_feature_column],  
    n_classes=2,  
  
    optimizer=tf.train.AdagradOptimizer(learning_rate=0.003))
```

Classifier

Predictions
(positive, negative)

TensorFlow NLP pretrained embeddings

TensorFlow Hub

text

Filter by

Language

Network

Publisher

Dataset

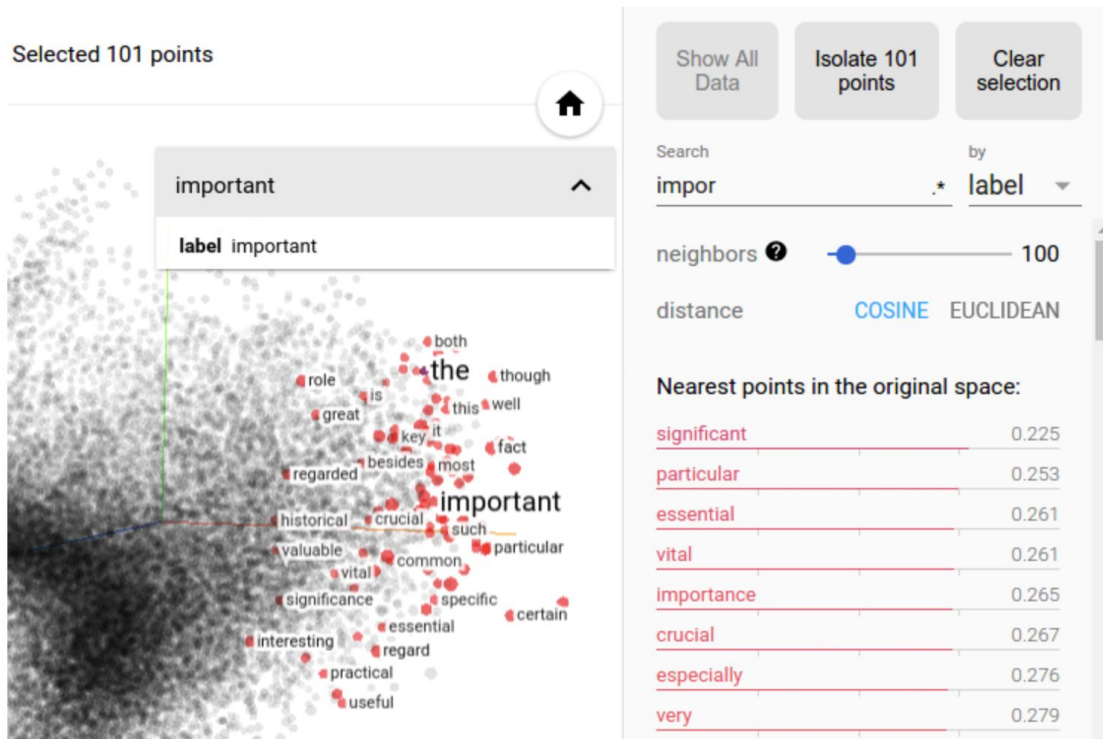
Module type

| Model Name | By | Text Embedding | Language | Description |
|----------------------------------|--------|----------------|--------------------------|---|
| universal-sentence-encoder | Google | text-embedding | DAN | Encoder of greater-than-word length text trained on a variety of data. |
| universal-sentence-encoder-large | Google | text-embedding | Transformer | Encoder of greater-than-word length text trained on a variety of data. |
| elmo | Google | text-embedding | 1 Billion Word Benchmark | Embeddings from a language model trained on the 1 Billion Word Benchmark. |
| nnlm-de-dim50-with-normalization | Google | text-embedding | Google News | Token based text embedding trained on German Google News 30B corpus. |
| nnlm-en-dim128 | Google | text-embedding | Google News | Token based text embedding trained on English Google News 200B corpus. |
| universal-sentence-encoder-lite | Google | text-embedding | Transformer | Encoder of greater-than-word length text trained on a variety of data. |

Word and sentence (!) level embeddings, trained on various corpora, including Google's gigantic

...more, not shown here

TensorBoard NLP visualizations



Recap

TensorFlow + NLP recap

- It's a generic, high-performing computation library + trained models + tools
- Directly applicable to natural language processing
- For practitioners
 - It has many useful pretrained models in TensorFlow Hub
 - It has a visual inspection and debugging tool
 - It supports multiple platforms (Unix, MacOS, Windows, iOS, Android)
- For researchers
 - It's a platform to create your own network architectures and optimization functions
 - Easy access to benchmarks (pretrained models)
 - Collect performance and behavior data with Tensorboard
- For both
 - Active community, many tutorials and books
 - High performance out of the box

More information

- [TensorFlow](#), [TensorFlow Hub](#), [TensorBoard](#)
- [TensorFlow graphs](#)
- [“Hello, TensorFlow!”](#) tutorial

Where to get started with NLP

- A great, easy to follow [text classification with movie reviews](#)