ParallelDots

Making a Contextual Recommendation Engine using Python and Deep Learning

By
Muktabh Srivastava
Co-founder ParallelDots

@muktabh Muktabh-Mayank

www.paralleldots.com

About ParallelDots

‘A technology company working on building state of the art artificial intelligence technology and giving easy access to anyone and everyone who needs it through our easy to use APIs’

Semantic Similarity

Analyze how close two sentences are with respect to each other

Text Classifier

Automatically classify the text according to your categories

Entity Extraction

Recognize the entities in the text to get the better picture of it

Sentiment Analysis

Determine the sentiment of the text that you are working on

Contextual Recommendation Engine

Need of a new recommendation engine

Pre-existing Solutions

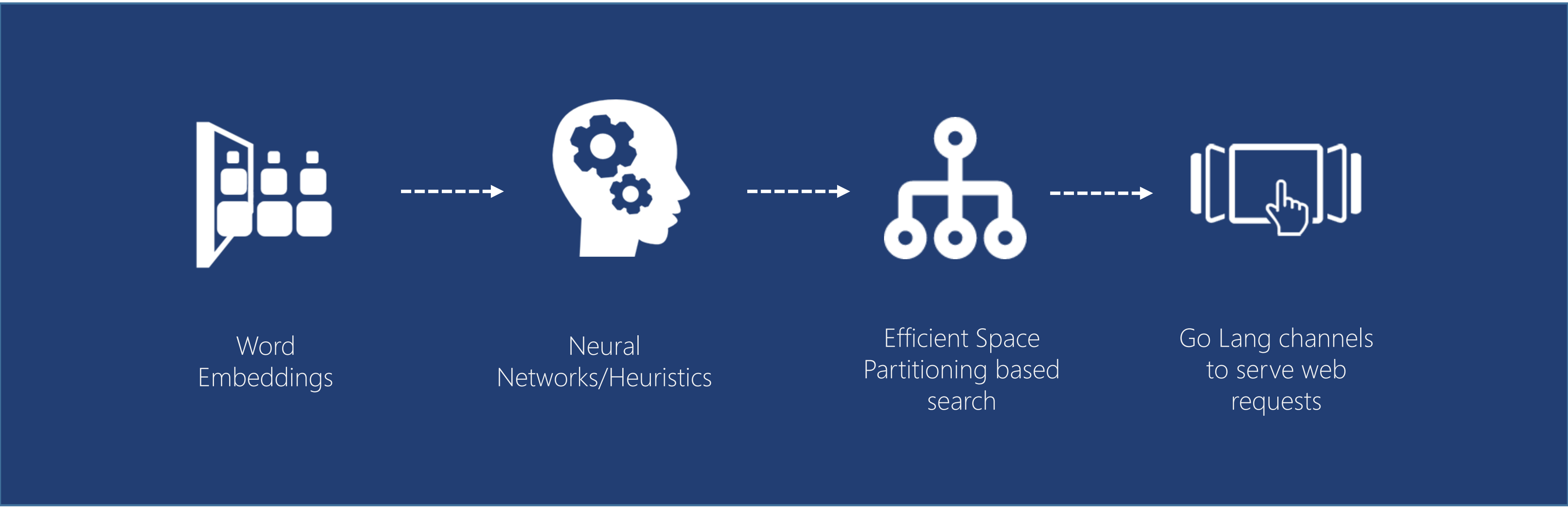
- Related Posts” plugin in CMS like Wordpress
- TFIDF based search which fails frequently
- Article tags can result in garbage results

Aims of new solution

- Should be more accurate than TFIDF tag search.
- Should be able to generate related posts for all articles.
- Should be cheap to deploy. (Its still “related posts” at the end of it)

Technology

We have four different layers which are interconnected to throw most accurate results at very high traffic blogs.



Word Embeddings

Neural Networks/Heuristics

Efficient Space Partitioning based search

Go Lang channels to serve web requests

Word Embeddings

Word Embeddings are dense low dimensional representation for each word

Google Word2Vec

- Most Popular
- Auto encoder which tries to reconstruct the co-occurence matrix
- Recently patented by Google

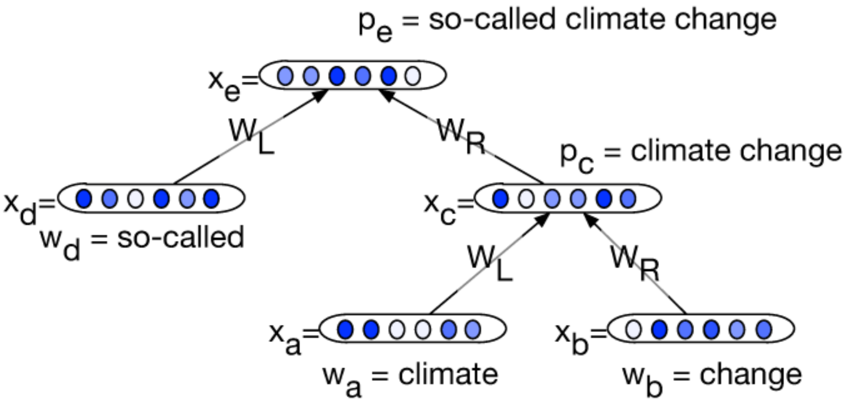
ParallelDots Tyrion

- Uses much lower RAM
- Can be trained in slower CPU
- Can be used on very large datasets with ease
- Open sourced

Deep Learning Models

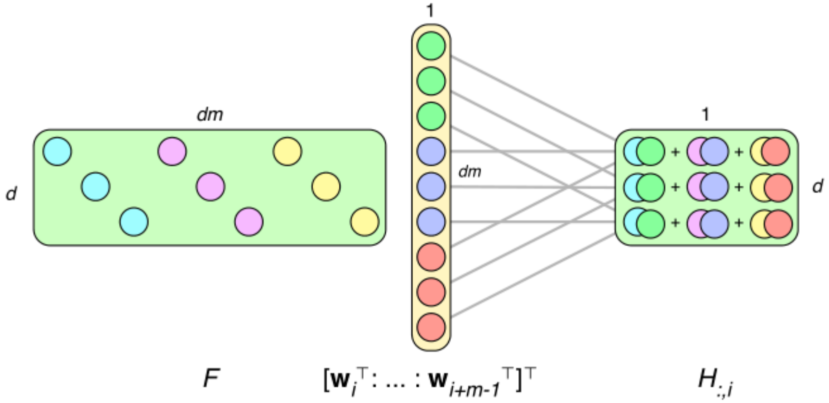
In Search Engine

- Recursive Neural Networks to combine Word Vectors into phrase Vectors for semantic closeness.
- Heuristically combining Word/Phrase Vectors for similar entity based near neighbors.



Others

- Convolutional Neural Networks to capture Sentiments in text.
- Recursive Neural Net based entity extraction.

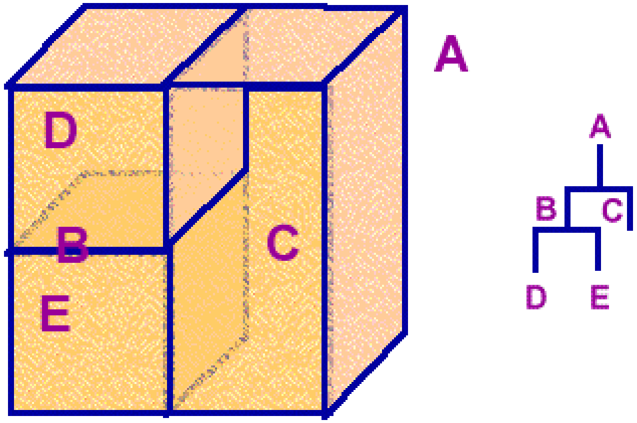


1
dm
d
dm
1
d
F
$[\mathbf{w}_i^\top : \ldots : \mathbf{w}_{i+m-1}^\top]^\top$
$H_{:,i}$

Vantage Point Tree

About VP Tree

- Arrange Document Representation on a Space partitioning Tree
- Uses VP Tree because of minimal requirements
- Makes query O(log n)< querytime < O(N)

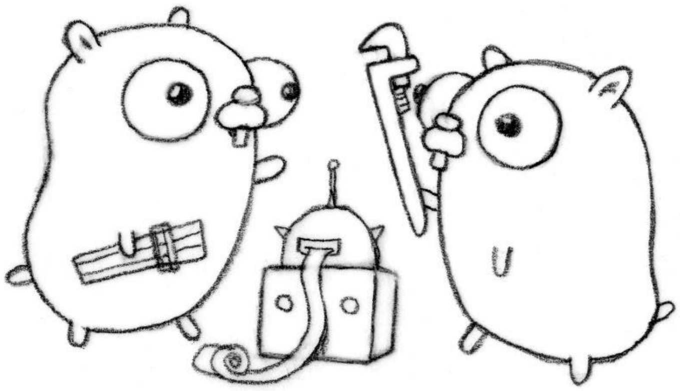


Our Implementation

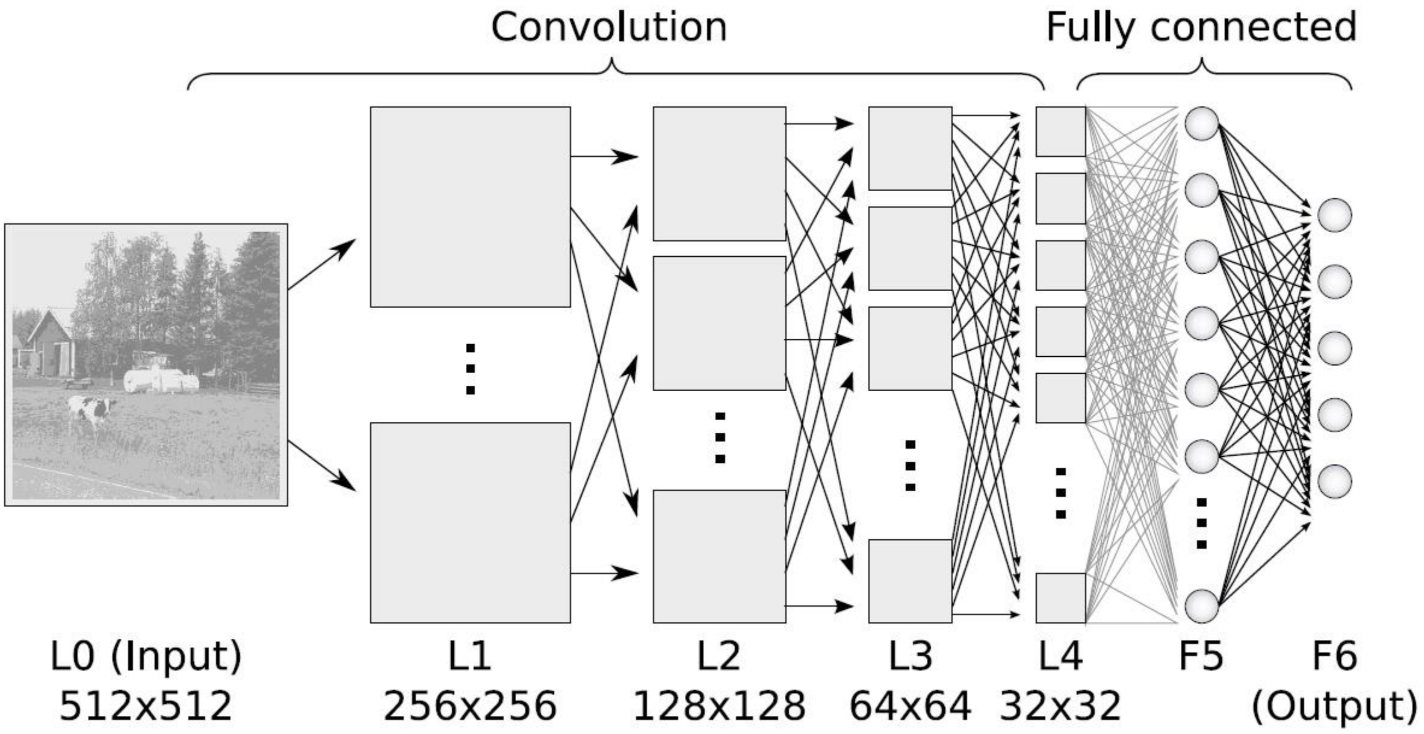
- Fast implementation in Numpy
- Right now documents divided into buckets and hosted one bucket/core using Python's multiprocessing
- Future Work: Shared memory model to make it work as true O(log(n))

Handling Traffic

- Analysis of incoming traffic using Pareto Principle
- Generation recommendations for viral unique articles
- Development Go-Lang channels to concurrently handle requests
- Golang channels to group these requests and deduplicate hits on Machine Learning infrastructure
- Combined with a caching layer (Redis) we handle up to 5000 concurrent users on a single box.



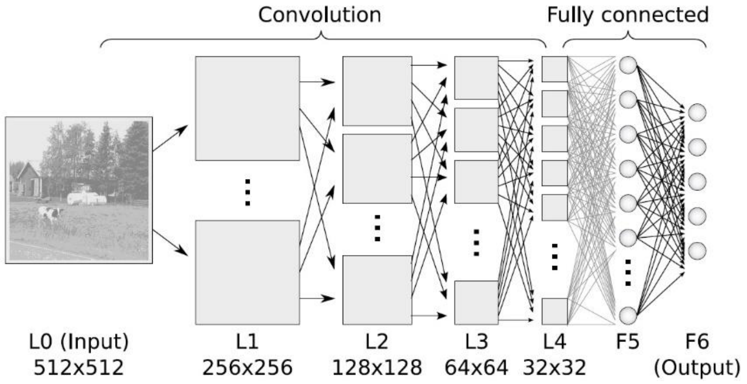
Basics of Deep Learning



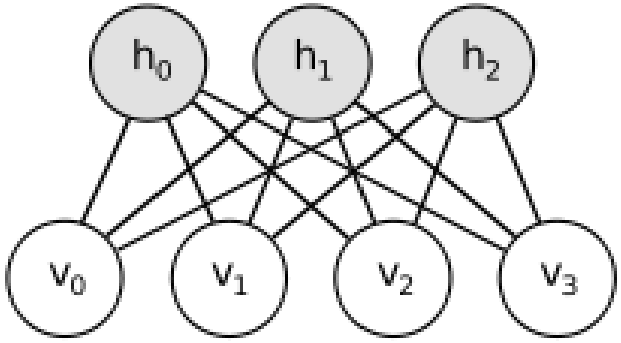
- Deep Learning is name given to multi layered Neural Networks.
- Layer(s) of weight are stacked on top of each other separated by layers of activation functions. Activation functions bring non-linearity into the learning, else multiple layers of weights would be same as one layer.
- They are trained by back propagation of errors generally by Gradient Descent.

Deep Learning continued…

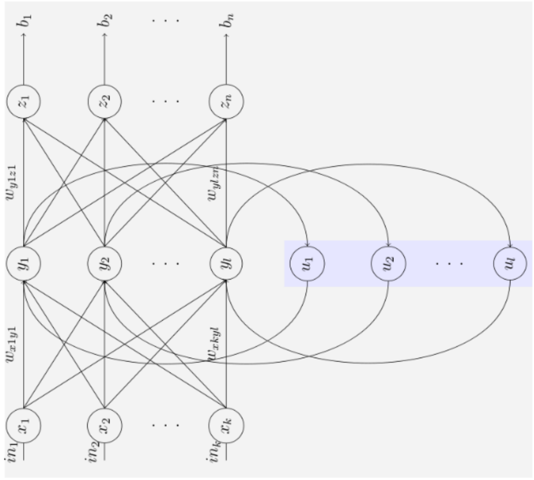
Types of Neural Networks



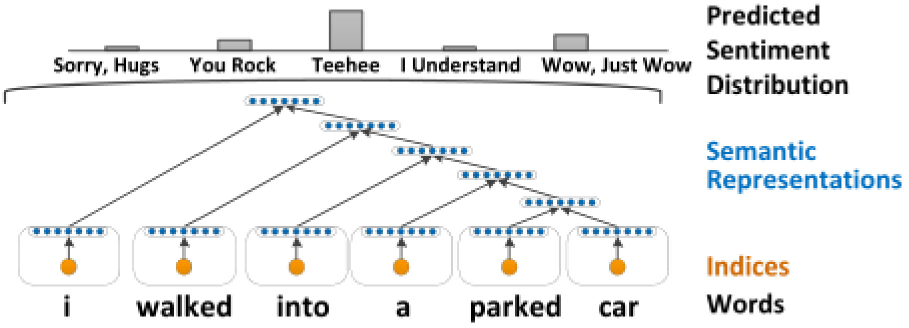
Convolutional



Boltzmann Machine(popular as RBMs)



Recurrent



Recursive

Images taken from Wikipedia, Socher et al(2012), deeplearningtutorials.net and University of Bonn slides.

Deep Learning continued…

Some vocabulary

Architecture

- Recurrent
- Recursive
- Convo NN
- DBN

Neural Net Units

- RBM
- Sigmoid
- Dropout
- ReLU
- LSTM
- tanh

Optimization

- Clipping
- Hessian
- MSGD
- Simulated Annealing
- SGD
- Momentum
- Adadelta
- Adagrad
- RMSprop

Implementing models

theano

Production Workhorse @ ParallelDots

Light CPU tryouts

Pylearn2

Kayak: Library for Deep Neural Networks

Lasagne

Lasagne is a lightweight library to build and train neural networks in Theano.

torch

Caffe

{P} cuda-convnet2

Fast convolutional neural networks in C++/CUDA

THANK YOU!

www.letsventure.com/paralleldots
www.angel.co/paralleldots

Like us or Follow us!