INDECOMM GLOBAL SERVICES
Partnerships@work

Technology | Learning | Platforms

Indecomm Technology

DIGITAL ENGINEERING AND ENGAGEMENT

Technology trends for 2018
- AI Foundation

Technology Trends 2018 – Gartner Research

INDECOMM GLOBAL SERVICES
Partnerships@work

Intelligent	Digital	Mesh
		
AI Foundation	Digital Twins	Blockchain
Intelligent Apps and Analytics	Cloud to the Edge	Event Driven
Intelligent Things	Conversational Platforms	Continuous Adaptive Risk and Trust
	Immersive Experience	

Agenda

INDECOMM GLOBAL SERVICES
Partnerships@work

- AI Growth Trends
- Artificial Intelligence – What is it?
- AI – Disciplines
- Definitions
- Case Studies

Artificial Intelligence - Growth

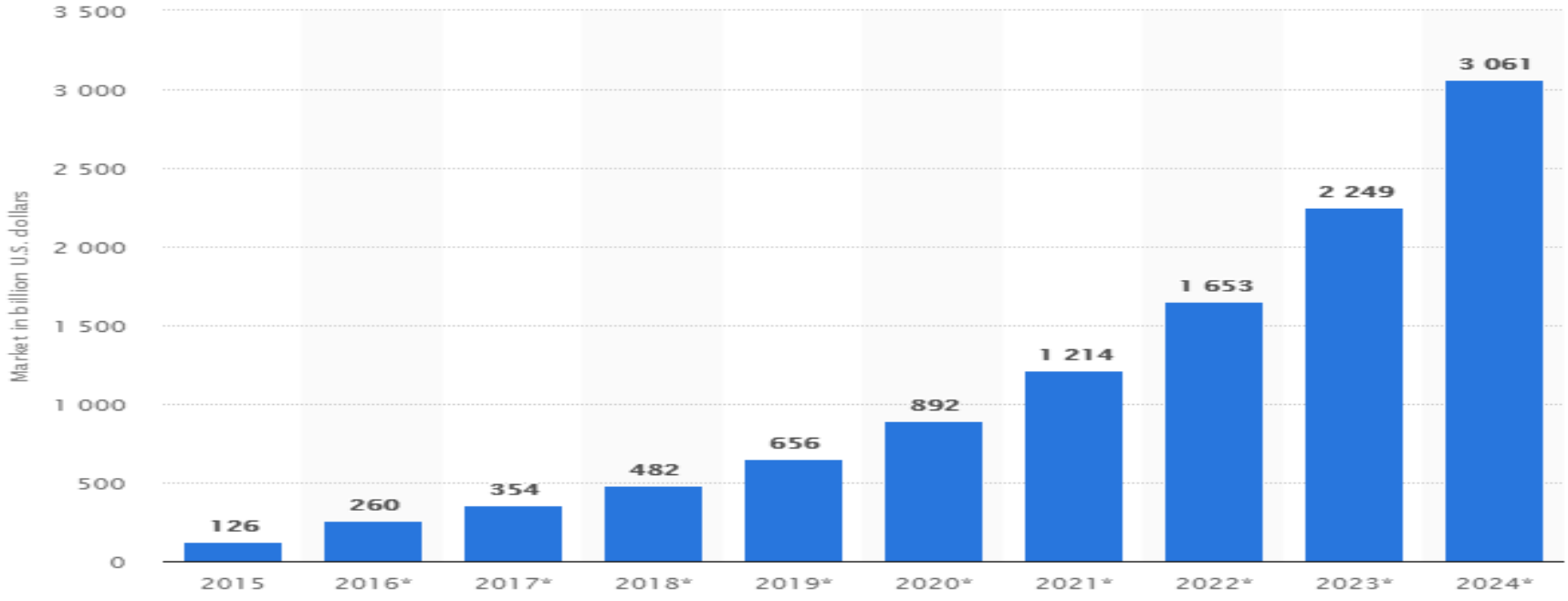
INDECOMM GLOBAL SERVICES
Partnerships@work

According to a study by Transparency Market Research:

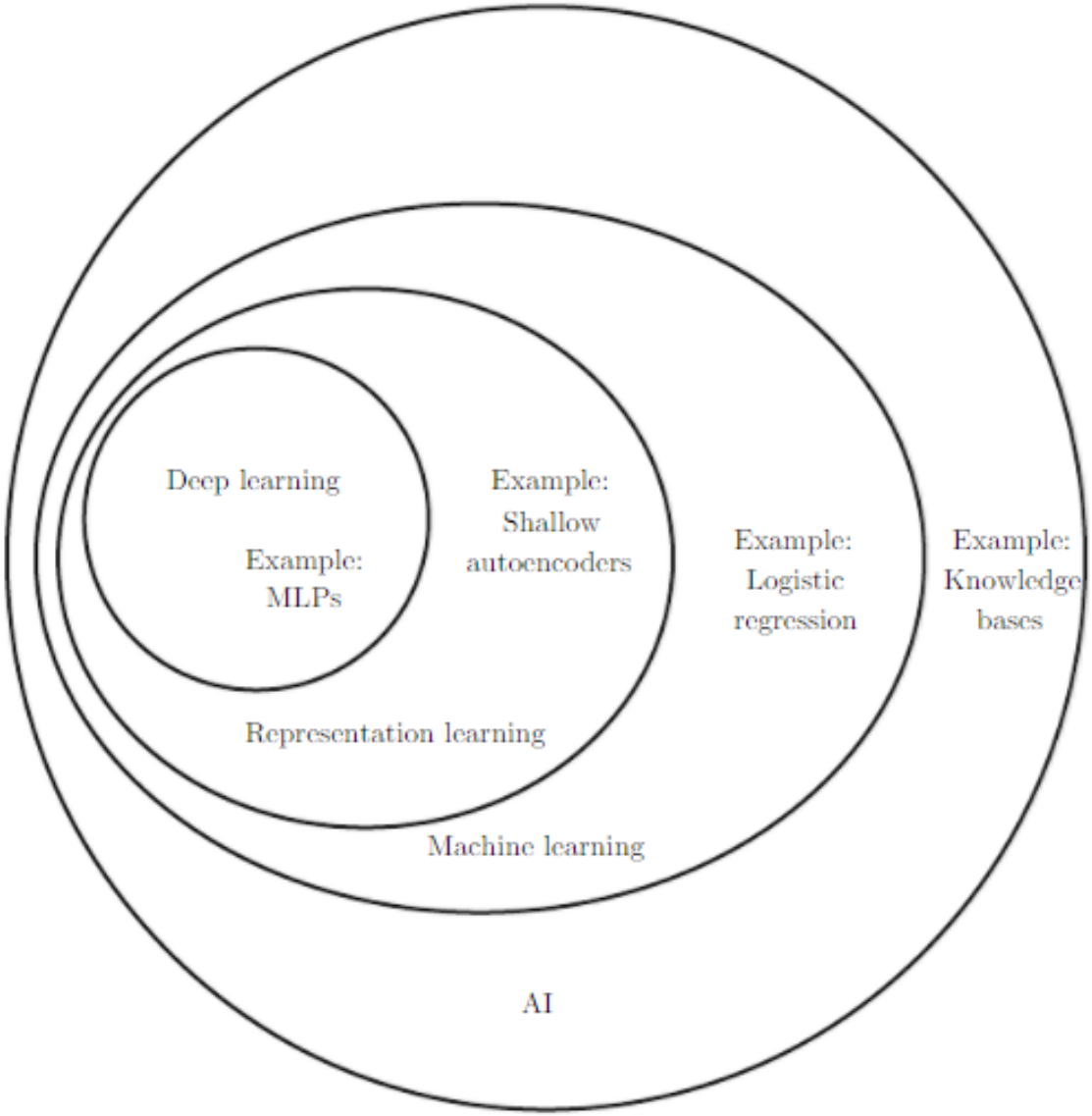
- Artificial Intelligence:
  - Compounded Annual Growth Rate (CAGR) of 36.1% between 2016 & 2024
  - To be valued at $3061 bn (2024) from $126 bn in 2015

- Machine Learning:
  - Compounded Annual Growth Rate (CAGR) of 38.4% between 2016 & 2024
  - To be valued at $19.86 bn (2024) from $1.07 bn in 2016

- AI industry has received more than $11.5 billion of investments in the last three years (till 2016) and is expected to raise over $6 billion of VC investments in 2017.

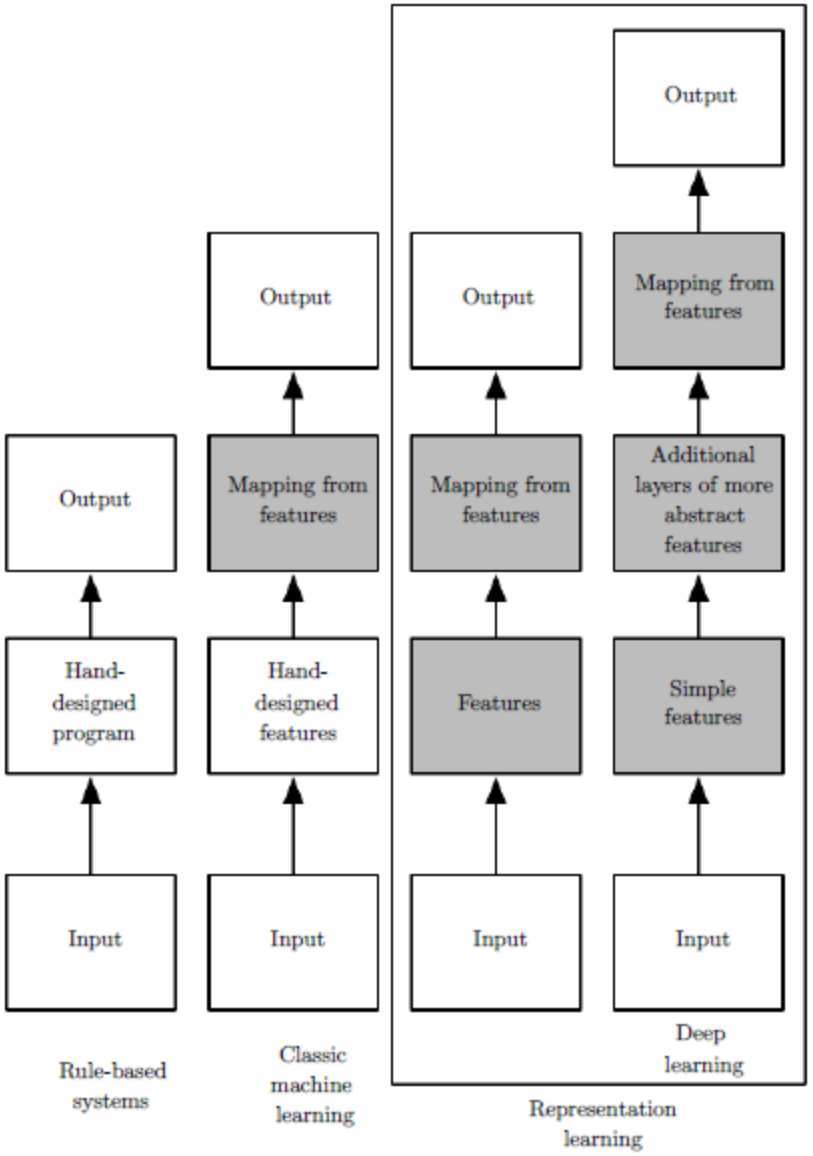


Artificial Intelligence



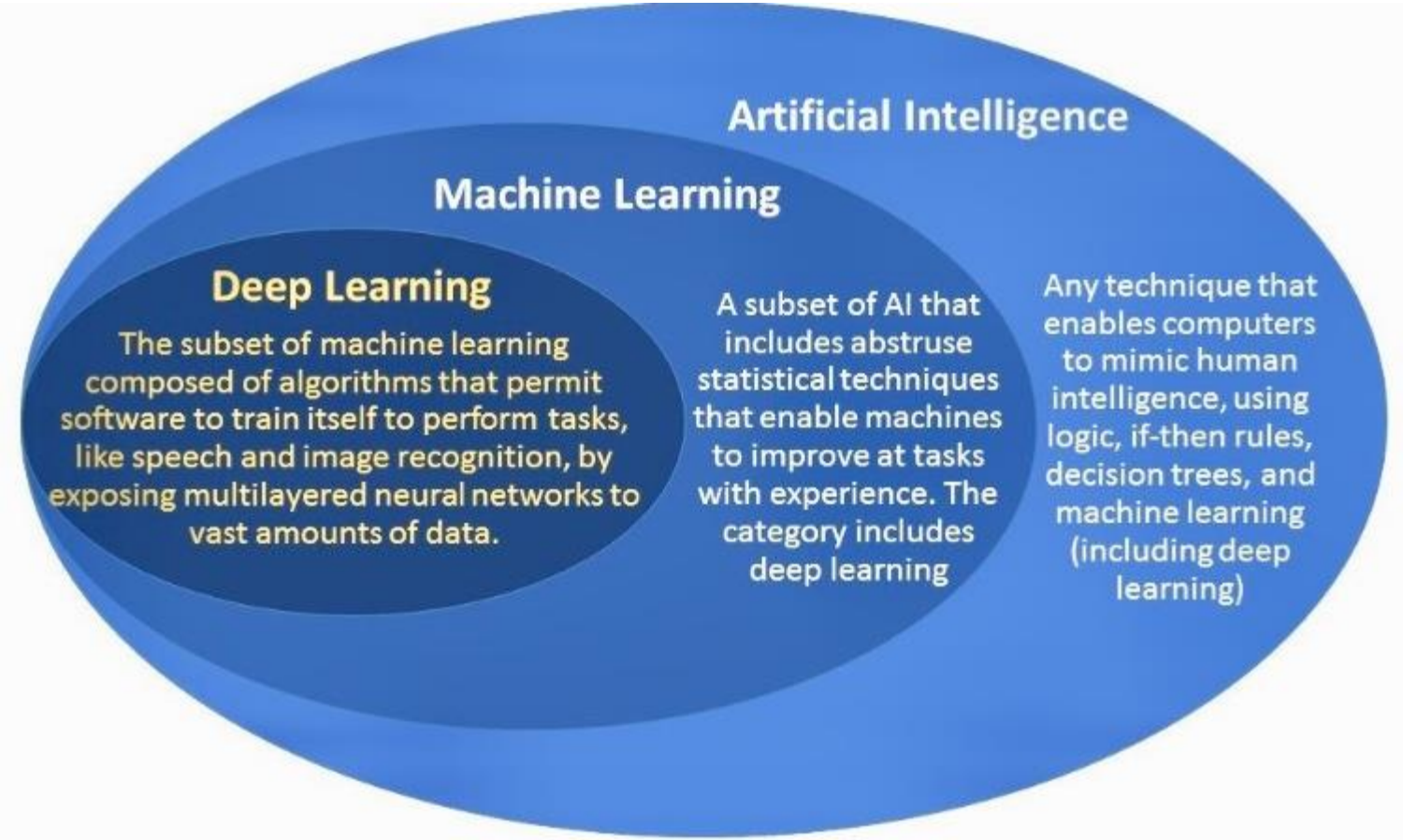
A Venn diagram showing how all the buzz words connect: Deep Learning & Machine Learning with AI

AI Disciplines

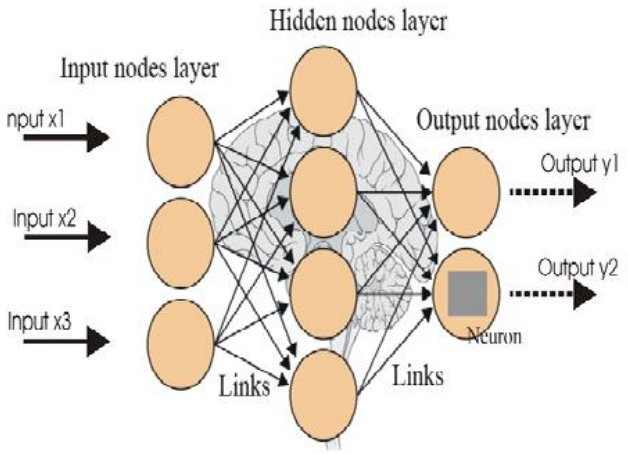


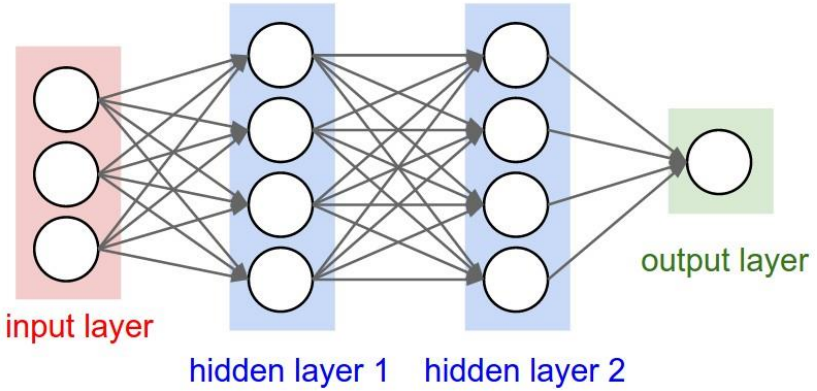
Artificial Intelligence

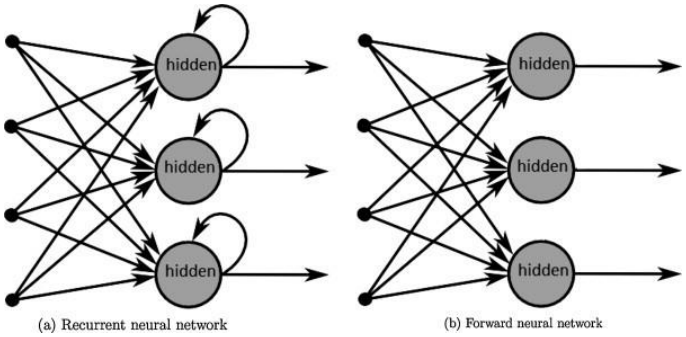
INDECOMM GLOBAL SERVICES
Partnerships@work

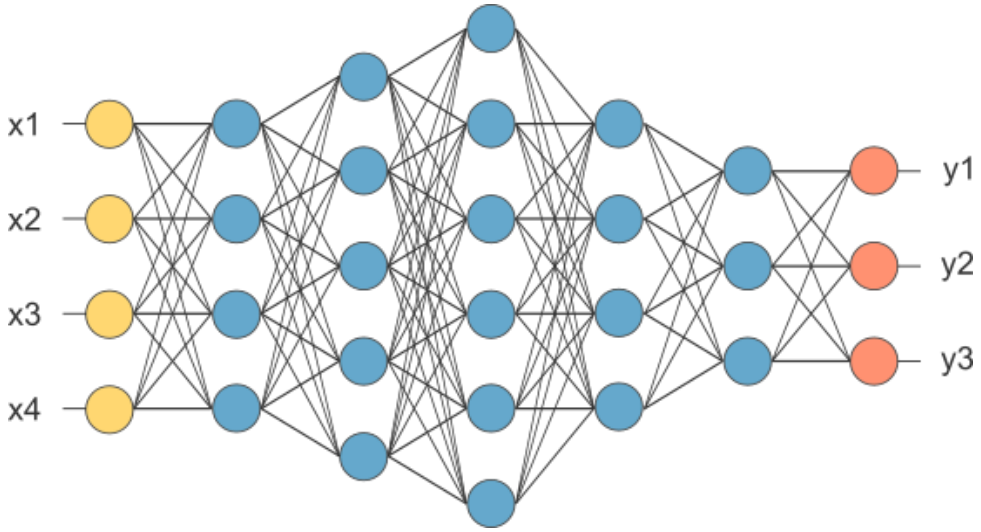


Neural Networks









Artificial Intelligence

INDECOMM GLOBAL SERVICES
Partnerships@work

Artificial Intelligence, essentially refers to the ability of a high-tech or computerized device to think like a human and possess/mimic human intelligence. AI effectively bestow what appears to be logic, memory and decision-making on these devices.
E.g. IoT, eCommerce, Siri, Google Now, Maps

Machine learning, on the other hand, is a specific type of AI that allows a piece of software to learn on its own. It explores the study and construction of algorithms that can learn from and make predictions based on the input data (training).
E.g. Numerical Computation, Linear Algebra, Probability, Algorithms

Deep learning, is a kind of machine learning, based on learning data representations, as opposed to task-specific algorithms. It provides a set of techniques to learn the features (arbitrary functions) and use them to perform a specific task, where some of these features can be adjusted to better predict the final outcome.
E.g. Artificial Neural Networks (Convolutional, Recurrent), Multi Layer Perceptron (MLP)

Machine Learning – Address Match

INDECOMM GLOBAL SERVICES
Partnerships@work

- **Address Match** between MLS Listing Table and Public Records, so as to improve record linkages
- Use a combination of various fields like Assessor's Parcel Number (APN), Unit #, Apt Name, Street Number, Street Name, County, City, Zip_Code, house attributes, neighborhood, etc. to determine a right match
- Usage of regular expressions to improve the data quality, and finding the right match
- Used fuzzy match algorithms to determine a score and identify the matches based on the ratio
- Usage of Soundex / Similar Sound, Geo Parsers, Levenshtein Distance algorithms, etc.

- Used **Deduplication** library (python):
  - **Dedupe** is a python based library that uses machine learning to perform de-duplication and entity resolution quickly on structured data.
  - Customized it to use the library to remove duplicate entries from MLS (based on fields like APN, US address, ZIP code, neighborhood, etc.)
  - The system is trained to find duplicates based on certain algorithms, even if all the data fields are not populated.

Machine Learning – Address Match

INDECOMM GLOBAL SERVICES
Partnerships@work

Score Distribution	Percentage of total unmatched listings	Comments
Scores 0 or NULL	72%	Listings with no matching PRs (Data not available)
Between 0 and 1	1.7%	False matches
Between 1 and 10	0.6%	Most of the matches are true matches (10% of the same may be false matches)
Between 10 and 50	1.5%	Most of the matches are true matches (3-5% of the same may be false matches)
Between 50 and 80	1.2%	True matches (A few attributes may be missing in MLS or PR hence reduced score)
Between 80 and 90	5%	True matches (A few attributes may be missing in MLS or PR hence reduced score)
Between 90 and 100	18%	Accurate matches

NLP – Deep Learning Example

INDECOMM GLOBAL SERVICES
Partnerships@work

Background:

- NLP (Natural language processing) is part of AI that has to do with language (usually written).
- The current algorithm runs against the public_remarks and classify the properties under the following attributes: Remodeling, Energy, Style, View, Flooring, Special Conditions, Heating Cooling.
- The algorithm is run against public_remarks column of the mls_listing table.

Our Approach:

- Input Data:
  - Created a training data set based on the provided public_remarks
  - Created a dictionary set with certain keywords and tagging them to the attributes

- Used Tensor Flow text analysis algorithm, with the following classifier models:
  - Convolutional Neural Networks
  - Recurrent Neural Networks
  - Vector Representation of words

- Support for multi-label text support
- Tried with different ratio of training data vs test data

NLP – Deep Learning Example

INDECOMM GLOBAL SERVICES
Partnerships@work

Option 2:

- Used FastText algorithm: A library for efficient text classification and representation learning
- Use pre-trained word vector and train the Word Vector based on the data
- Usage of Frequency or Prediction based Word Embeddings. It uses entire Wikipedia dump to form word representations in a vector / matrix (Word2Vec)
- FastText is much faster as compare to TensorFlow, and can tag / embed millions of words in a matter of few hours.

Exercise

INDECOMM GLOBAL SERVICES
Partnerships@work

- Build a model to predict the salary of an offered candidate?

References

- https://www.statista.com/statistics/621035/worldwide-artificial-intelligence-market-revenue/
- "*Deep Learning*" by Ian Goodfellow, Yoshua & Aaron (http://www.deeplearningbook.org/)
- https://www.quora.com/What-is-the-difference-between-artificial-intelligence-machine-learning-and-deep-learning-Is-deep-learning-a-part-of-machine-learning
- https://www.pyimagesearch.com/2016/09/26/a-simple-neural-network-with-python-and-keras/
- http://www.opennn.net/