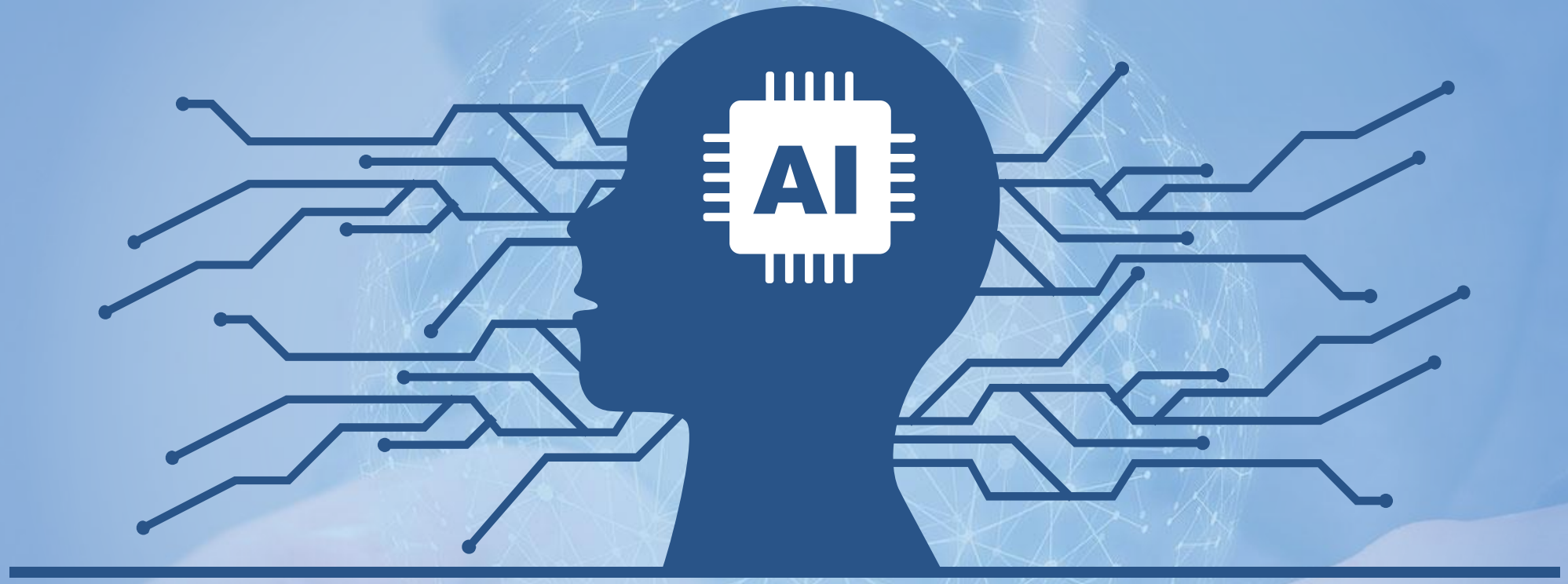


Welcome to Julie’s Portfolio

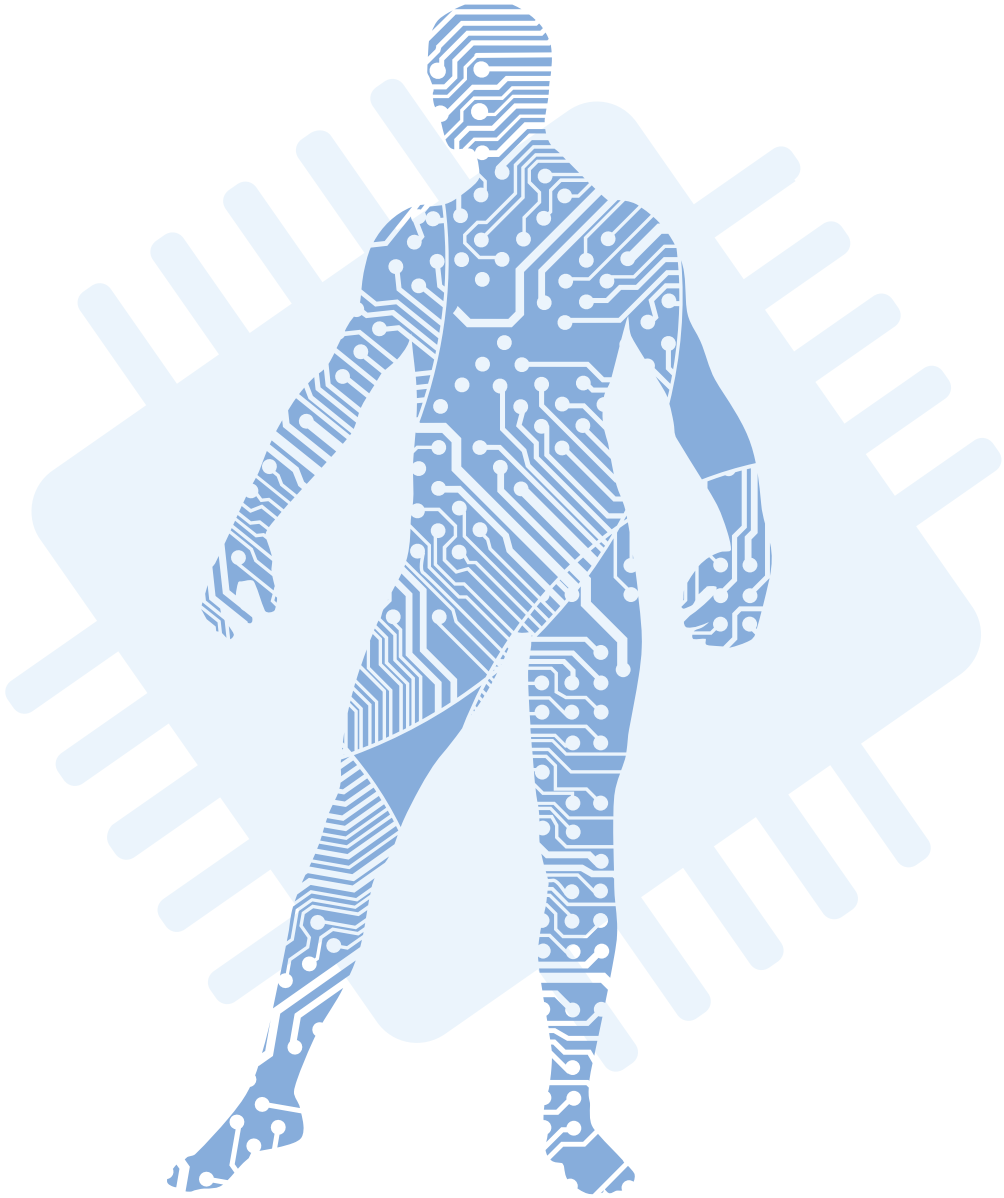
01 About Me
Self-introduction

02 Professional Experience
Background, experience, skills

03 Deep Learning Projects
Presentation of deep learning works

04 Machine Learning Projects
Presentation of machine learning works

05 Summary
Portfolio summary



About Me

Julie Yan

Machine Learning Engineer
QA Engineer
Sports Lover

In 2015, I joined MSC software as a QA engineer, when I graduate from University of Southern California, and got my master degree. In 2016, msc former CEO, Dominic Gallello encouraged everybody to study machine learning, which discipline gets more and more popular and powerful. Online courses are provided to all employees, and that was the start point of my journey in artificial intelligence.

Over the past three years, I have dedicated a considerable amount of time and effort to machine learning. Work hard to equip and prepare myself to be a skilled machine learning engineer. I have completed a number of courses in machine learning, deep learning, data science, statistics, probability, computer science, data structure and algorithm; as well as challenging projects on computer vision and natural language processing.

Time fly, hopefully my career in machine learning can take off too.

Experience

2009-2013

Beihang University

Aerospace Engineering
Bachelor of Science

2013-2015

University of Southern California

Mechanical Engineering
Master of Science

2015-current

MSC Software Corporation

QA Engineer

2016-future

Machine Learning

Start journey of artificial intelligence, machine learning, deep learning and data science

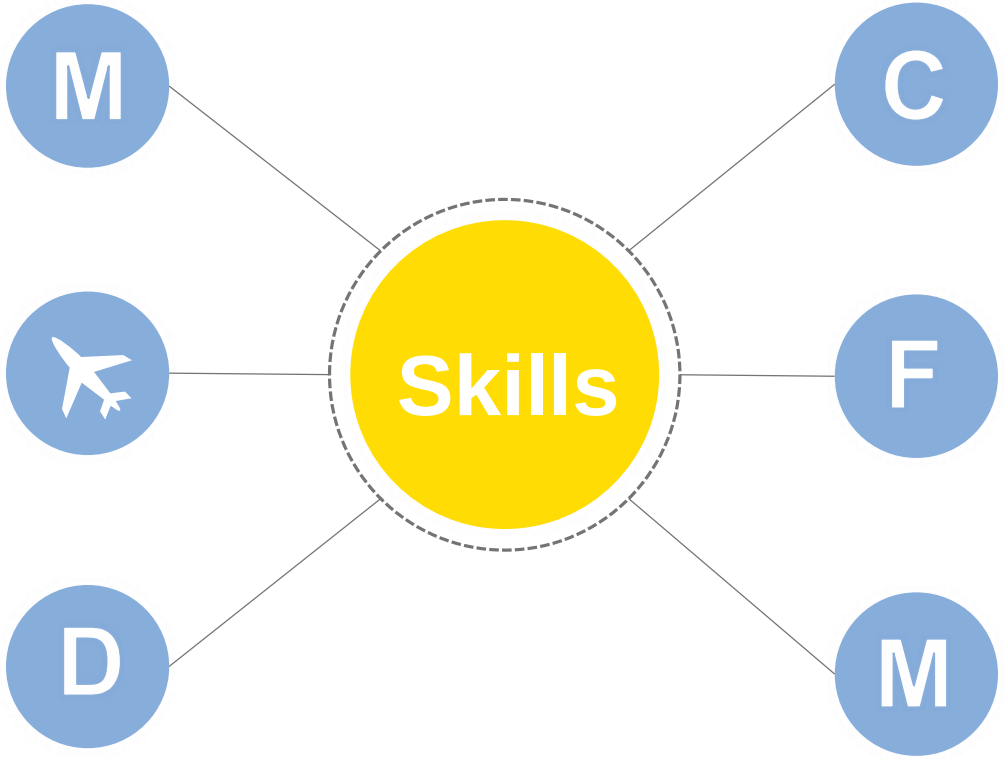
Skills

Machine Learning

- Linear Regression
- Logistic Regression
- Decision Tree
- KNN
- K-Means
- KD-tree
- Support Vector Machines
- Naïve Bayes
- Random Forest
- PCA
- Adaboost
- Gradient Boost
- XGBoost

Deep Learning

- CNN
- R-CNN
- Fast R-CNN
- Faster R-CNN
- Mask R-CNN
- YOLO
- VGG
- ResNet
- Unet
- Deconvolution Net
- RNN
- LSTM
- GRU
- Attention
- Bi-directional



Computer Science

- Data Structure
- Data Algorithm
- Python
- Windows / Linux

Framework

- Tensorflow
- Keras
- Scikit-learn
- Nltk
- Numpy
- Pandas
- Scipy
- Matplotlib
- Beautiful soup
- cv2

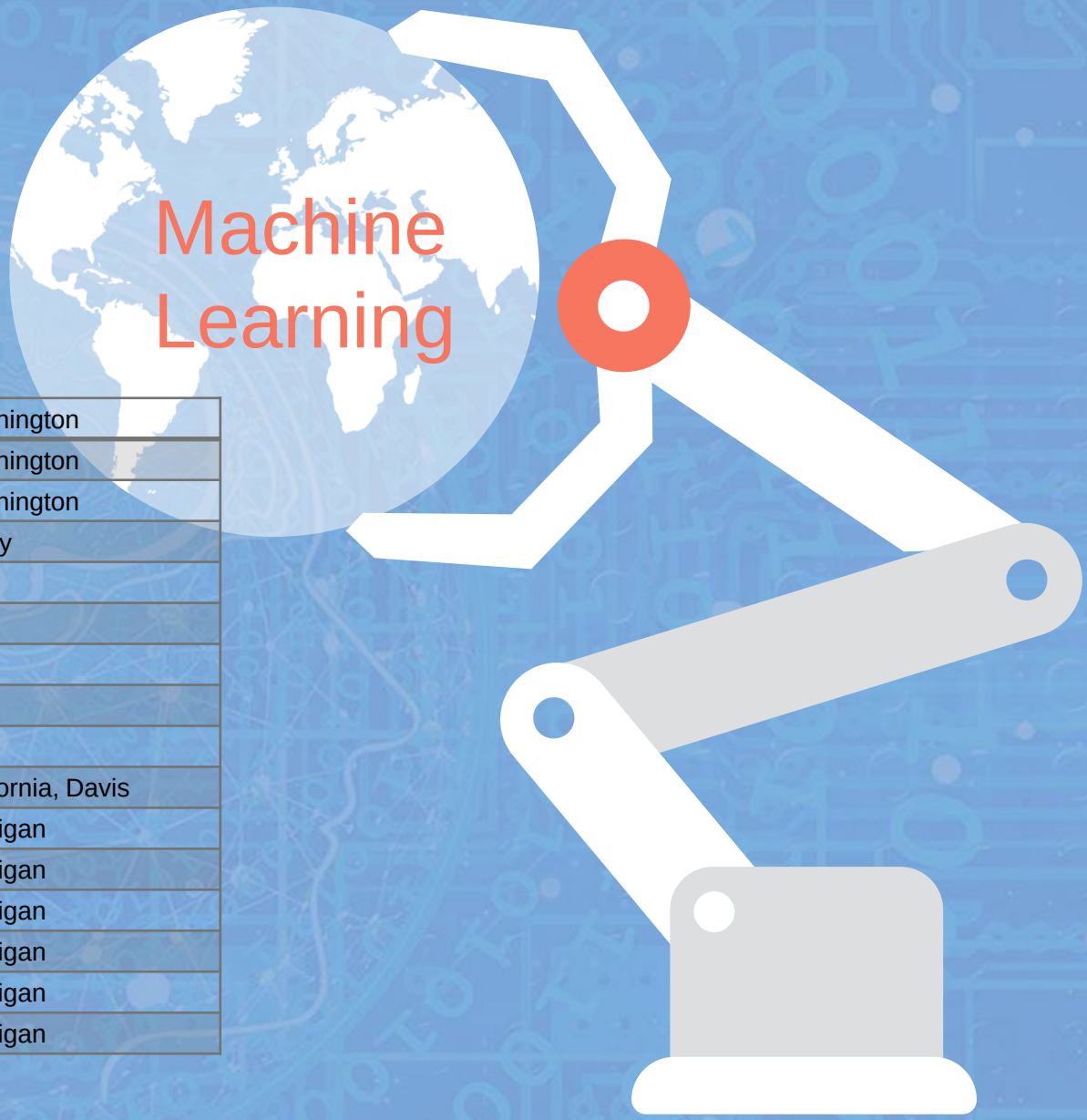
Mathematics

- Statistics
- Probability
- Linear Algebra
- Applied Mathematics
- Engineering Mathematics

Certificates

In the past three years, I have dedicated a considerable time and efforts on machine learning and deep learning. Here is a list of certificates I achieved for data science and machine learning.

· **Machine Learning Classification**	University of Washington
· **Machine Learning Regression**	University of Washington
· **Machine Learning Clustering & Retrieval**	University of Washington
· **Machine Learning**	Stanford University
· **Neural Networks and Deep Learning**	Deeplearning.AI
· **Structuring Machine Learning Projects**	Deeplearning.AI
· **Convolutional Neural Network CNN**	Deeplearning.AI
· **Recurrent Neural Network RNN**	Deeplearning.AI
· **Hyper parameter tuning, Regularization and Optimization**	Deeplearning.AI
· **SQL for Data Science**	University of California, Davis
· **Python Data Structures**	University of Michigan
· **Using Python to Access Web Data**	University of Michigan
· **Using Databases with Python**	University of Michigan
· **Retrieving, Processing, and Visualizing Data with Python**	University of Michigan
· **Data Science in Python**	University of Michigan
· **Applied Plotting, Charting & Data Representation in Python**	University of Michigan



Deep Learning Projects

DNN CNN RNN Project Presentation

Segmentation Project

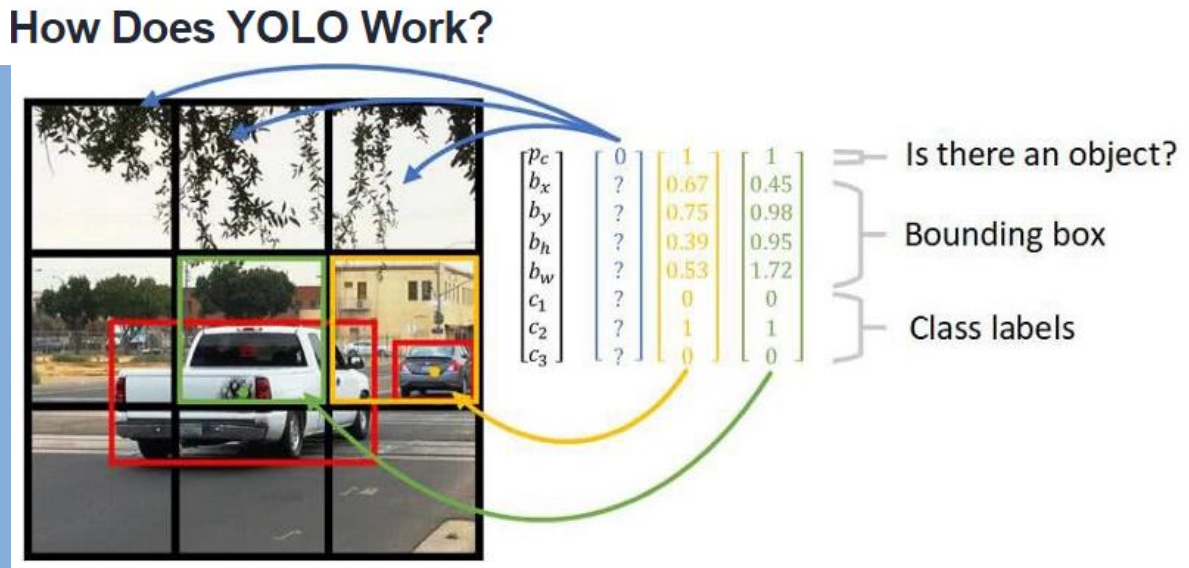
The task is to do both categorization and segmentation of rich and complete apparel attributes, an important step toward real-world applications.

- Project: Kaggle - iMaterialist (Fashion) 2019 at FGVC6, https://www.kaggle.com/c/imaterialist-fashion-2019-FGVC6
- Model Structure: **Mask RCNN**
- Code : https://github.com/JialieY/Kaggle-iMaterialist

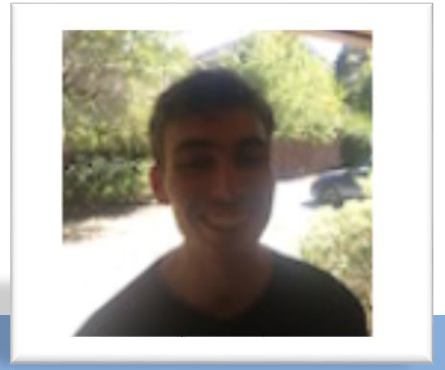
Object Detection Project

The task is to detect objects like traffic lights, truck, car, and pedestrian in the image and videos.

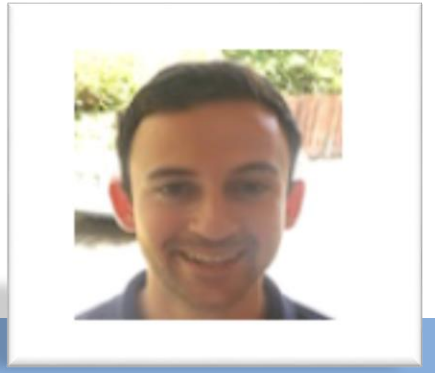
- Project: Autonomous driving Car detection
- Model Structure: **Yolo model**
- Code : https://github.com/JialieY/dl_ng/blob/master/Course4/Week3/Autonomous%20driving%20application%20-%20Car%20detection%20-%20v1.ipynb



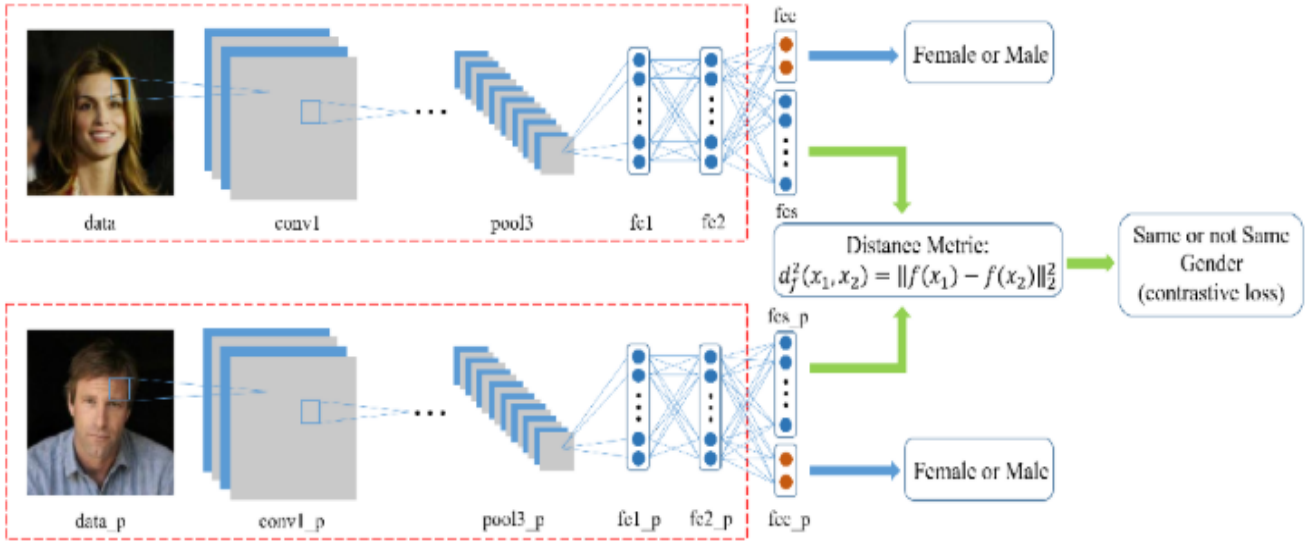
Face Recognition & Verification Project



It's not Kian, please go away
(0.86224014, False)



It's younes, welcome home!
(0.65939283, True)



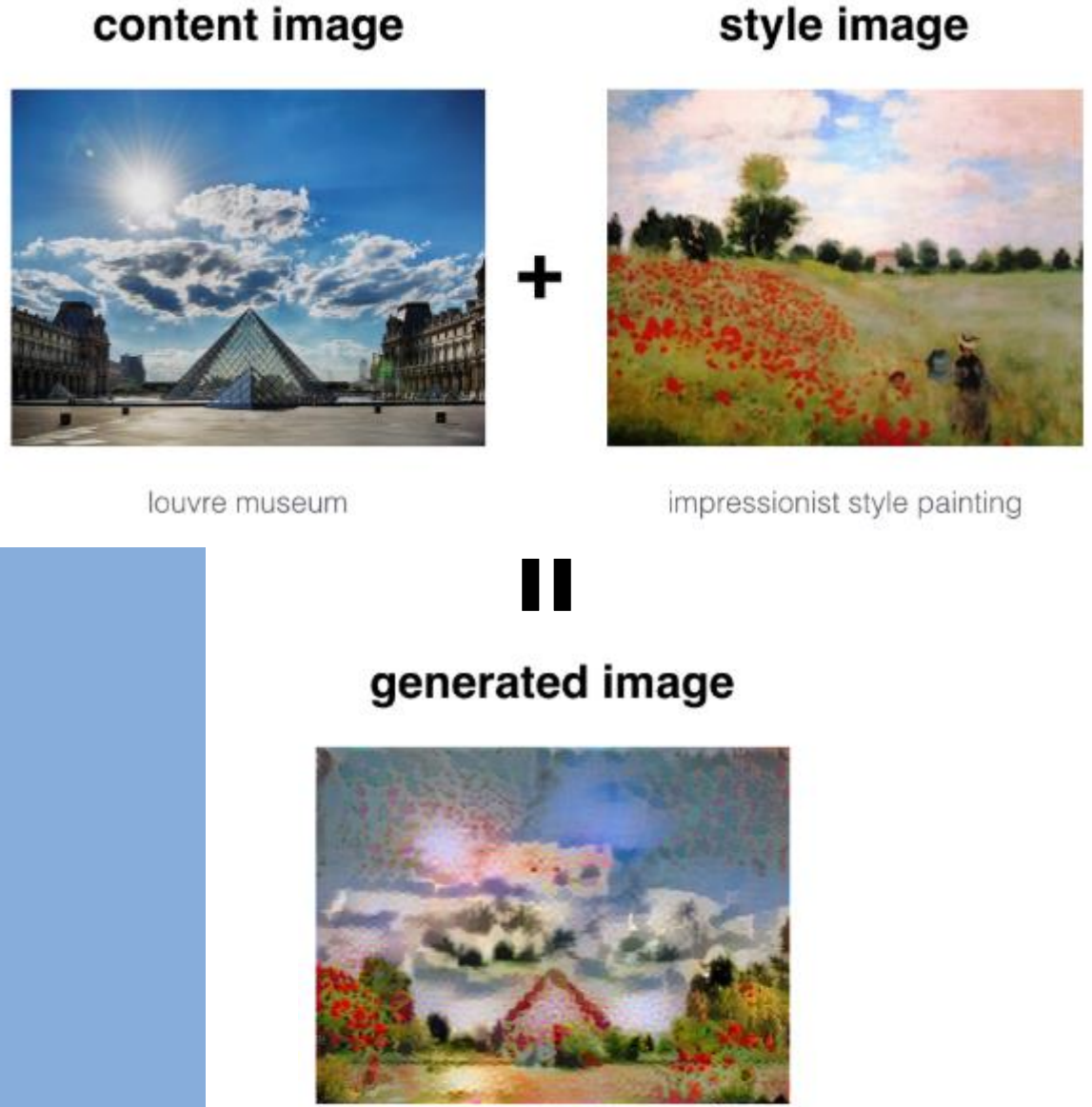
The task is to recognize and verify people with existing images in the database.

- Project: Face Recognition & Verification
- Model Structure: **FaceNet, DeepFace, Siamese network**
- Code : https://github.com/JialieY/dl_ng/blob/master/Course4/Week4/Face/Face%20Recognition%20for%20the%20Happy%20House%20-%20v2.ipynb

Neural Style Transfer Project

The task is to generate a new image with given content and given art style.

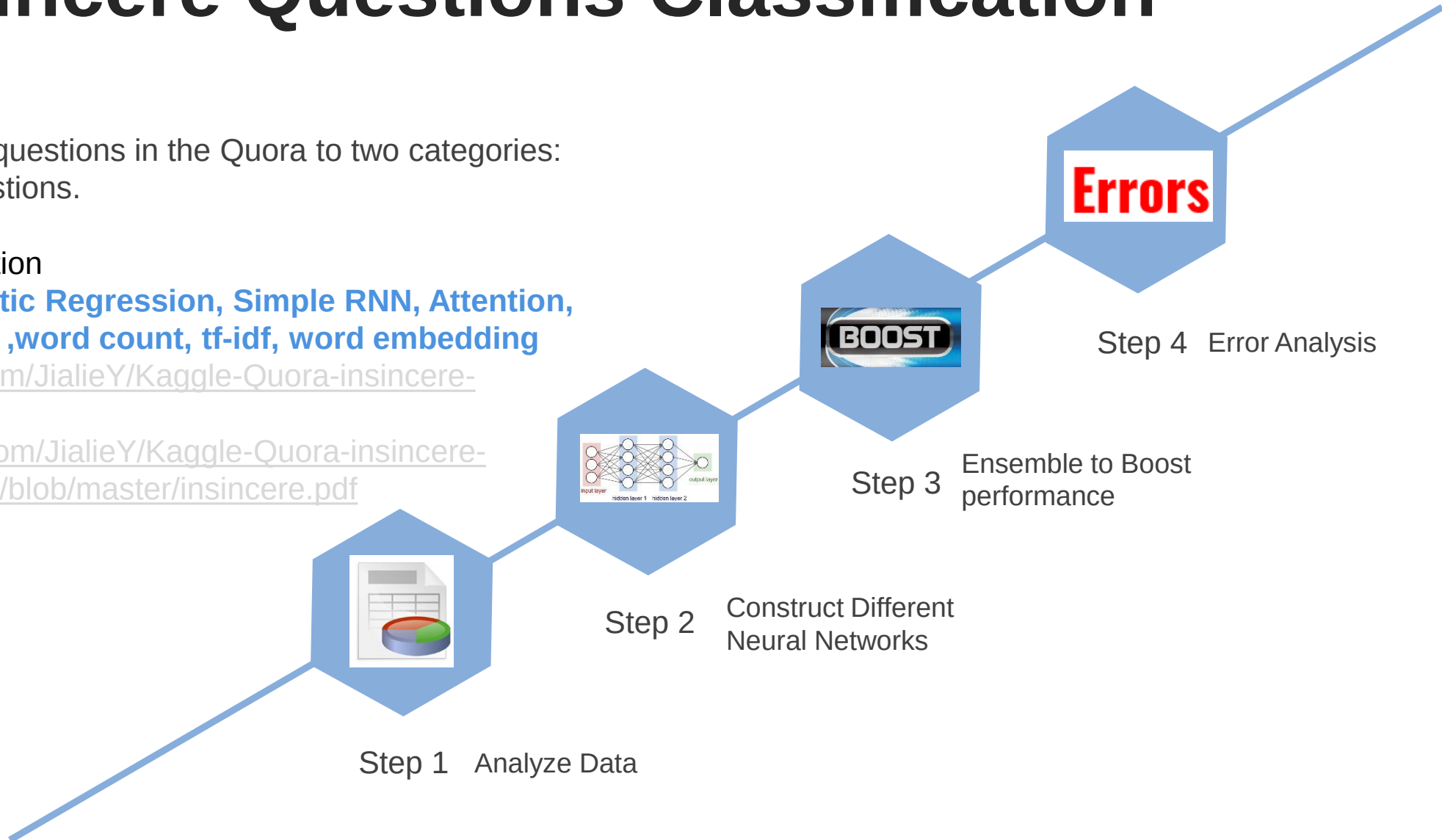
- Project: Art generation
- Model Structure: VGG16, VGG19
- Code : https://github.com/JialieY/dl_ng/blob/master/Course4/Week4/Neural%20Style%20Transfer/Art%20Generation%20with%20Neural%20Style%20Transfer%20-%20v1.ipynb

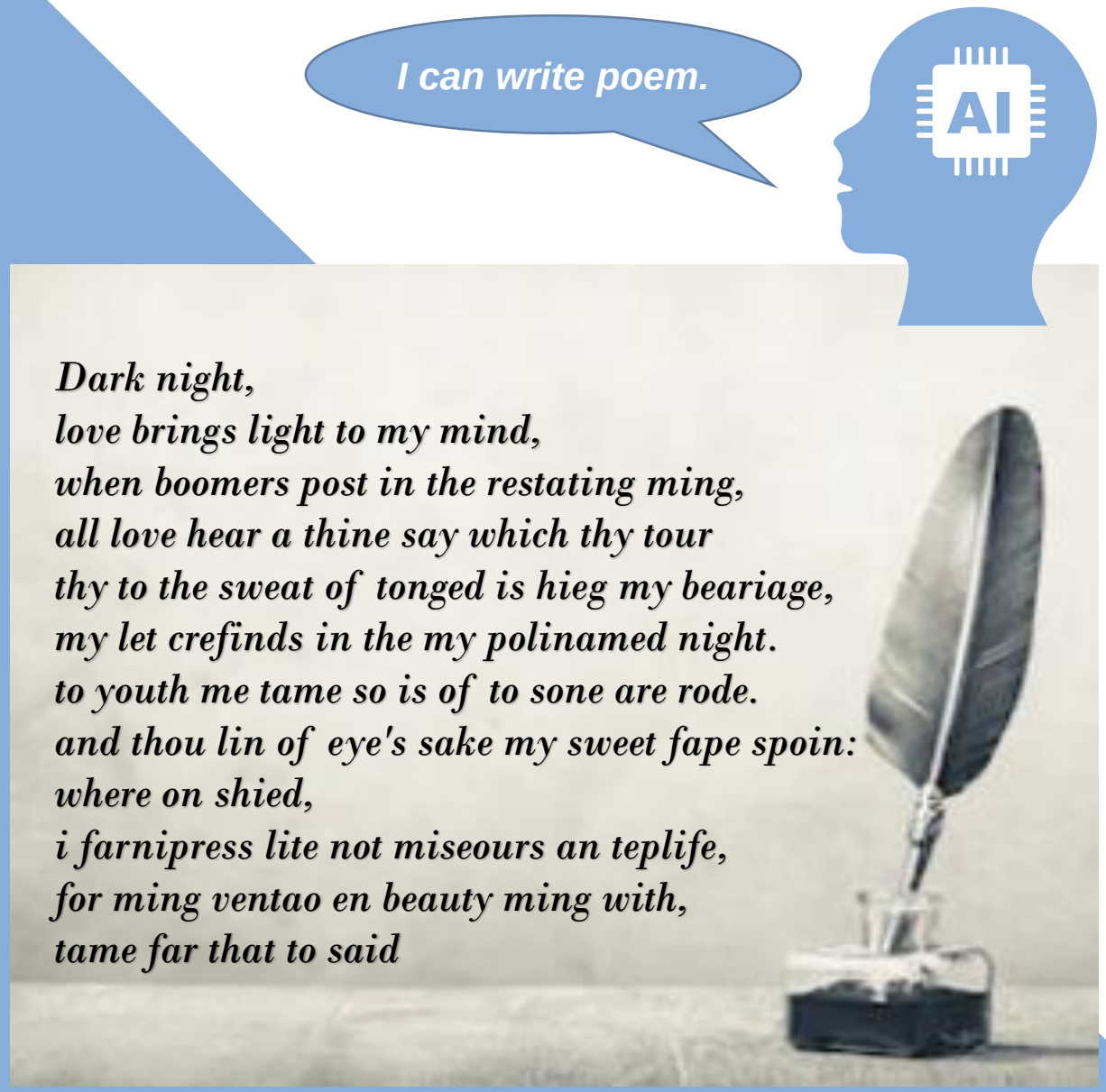


Quora Insincere Questions Classification

The task is to classify the questions in the Quora to two categories: sincere and insincere questions.

- Project: Text Classification
- Model Structure: **Logistic Regression, Simple RNN, Attention, LSTM, GRU, XGboost ,word count, tf-idf, word embedding**
- Code : https://github.com/JialieY/Kaggle-Quora-insincere-questions-classification
- Report: https://github.com/JialieY/Kaggle-Quora-insincere-questions-classification/blob/master/insincere.pdf





Sequence Generation Project

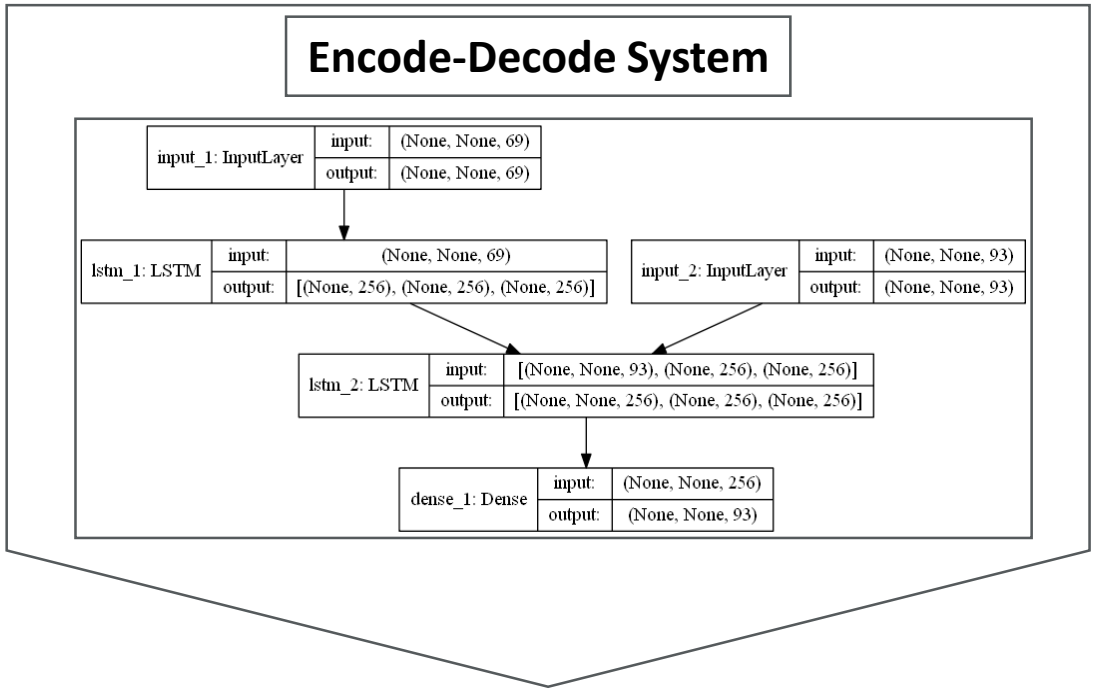
The technique is to learn the given pattern, and generate the similar output.

- Project: Name generation, Article generation, Music generation
- Model Structure: **RNN, LSTM, Attention**
- Code : https://github.com/JialieY/dl_ng/blob/master/Course5/Week1/Jazz https://github.com/JialieY/dl_ng/tree/master/Course5/Week1/Dinosaurus

Translator Project

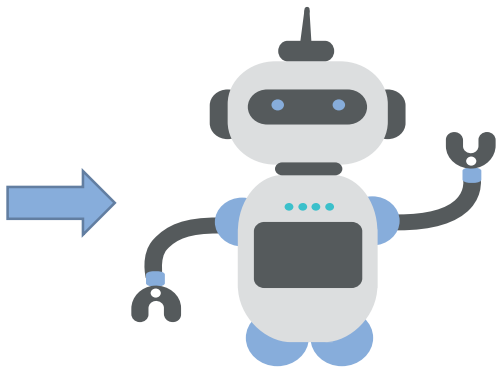
The task is to translate English sentence to French sentence.

- Project: English – French Translator
- Model Structure: **LSTM, Encode-Decode system**



English

Be calm.
Be cool.
Be fair.



French

Soyez calmes !
Sois détendu !
Sois juste !

Machine Learning Projects

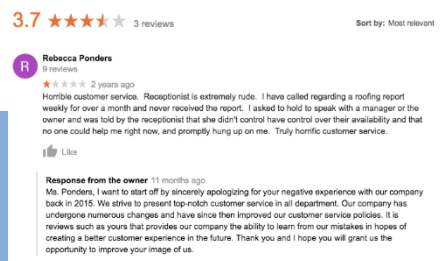
Supervised & unsupervised learning
Projects presentation

Amazon Review Classification

The task is to predict whether the sentiments about a product are positive or negative.

- Project: Amazon product reviews classification
- Data input: Reviews
- Model Structure: **Linear regression, Logistic regression, bag of words, tf-idf, L2 regularization, L1 regularization, Stochastic Gradient Descent**

Problem

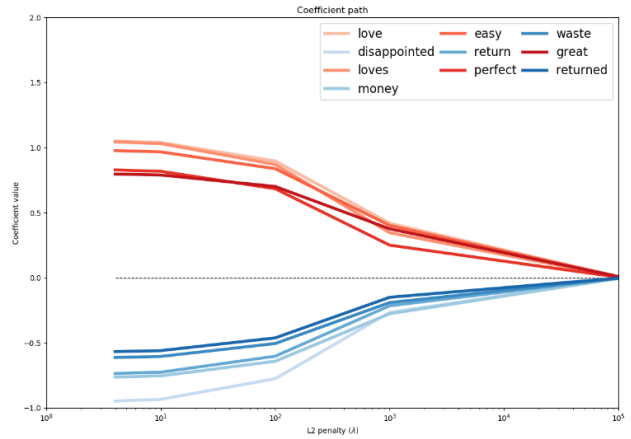


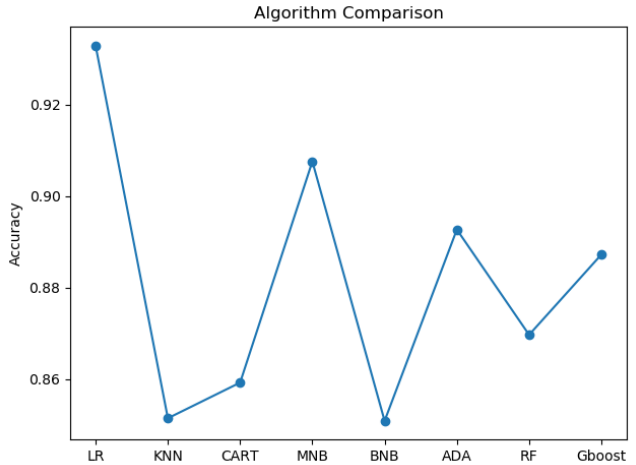
Find out it is

Input

review	rating	sentiment
All of my kids have cried non-stop when I tried to ween them off their pacifier, until I found Thumbuddy To Love	5	1
We wanted to get something to keep track of our child's milestones and this is a cute option. There aren't many	5	1
My daughter had her 1st baby over a year ago. She did receive and fill up a First Year Calendar. When her son was	5	1
One of baby's first and favorite books, and it is washable! I gave 1 less star than perfect because I'd like to see so	4	1
Very cute interactive book! My son loves this book! The bright colorful illustrations make this great for babies &	5	1
Beautiful book, I love it to record cherished times in my great granddaughters life with the beautiful pastel pink	5	1
Try this out for a spring project !Easy ,fun and affordable wall decals ...Fine quality and brightens up any room.. 5	5	1
very nice Divine Mercy Pendant of Jesus now on chain around my neck. Love It! Divine Mercy represents Jesusw	5	1
We bought the pins as my 6 year old Autistic son was able to open or break open normal safety pins we needed t	4	1
It has been many years since we needed diaper pins, but I do like to pin my socks together through the wash and	5	1
We found this book at a rummage sale and found it to be so useful, especially when several people are helping v	5	1

Algorithm





Loan Classification

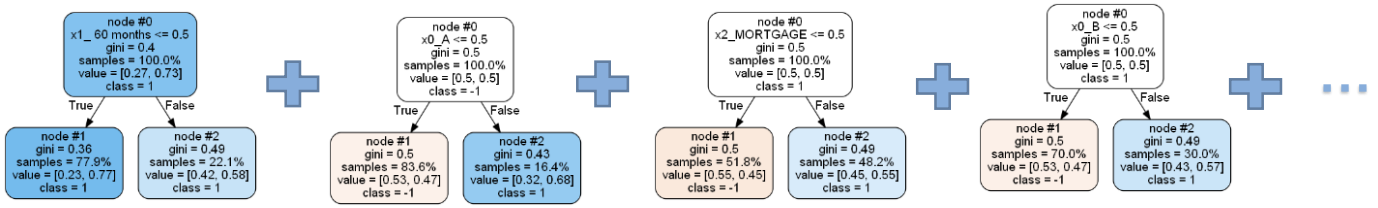
The task is to predict whether a loan is safe or not.

- Project: Loan data classification
- Data input: numbers, text, and etc.
- Model Structure: **Decision Tree, Adaboost**

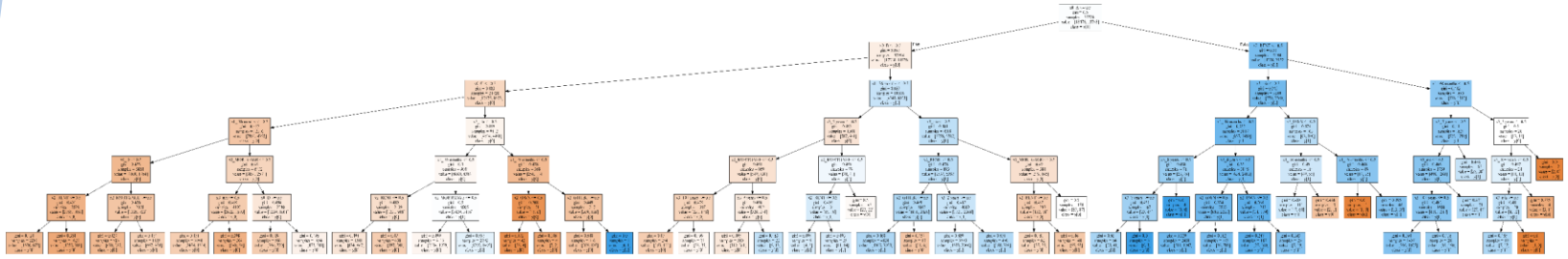
Input

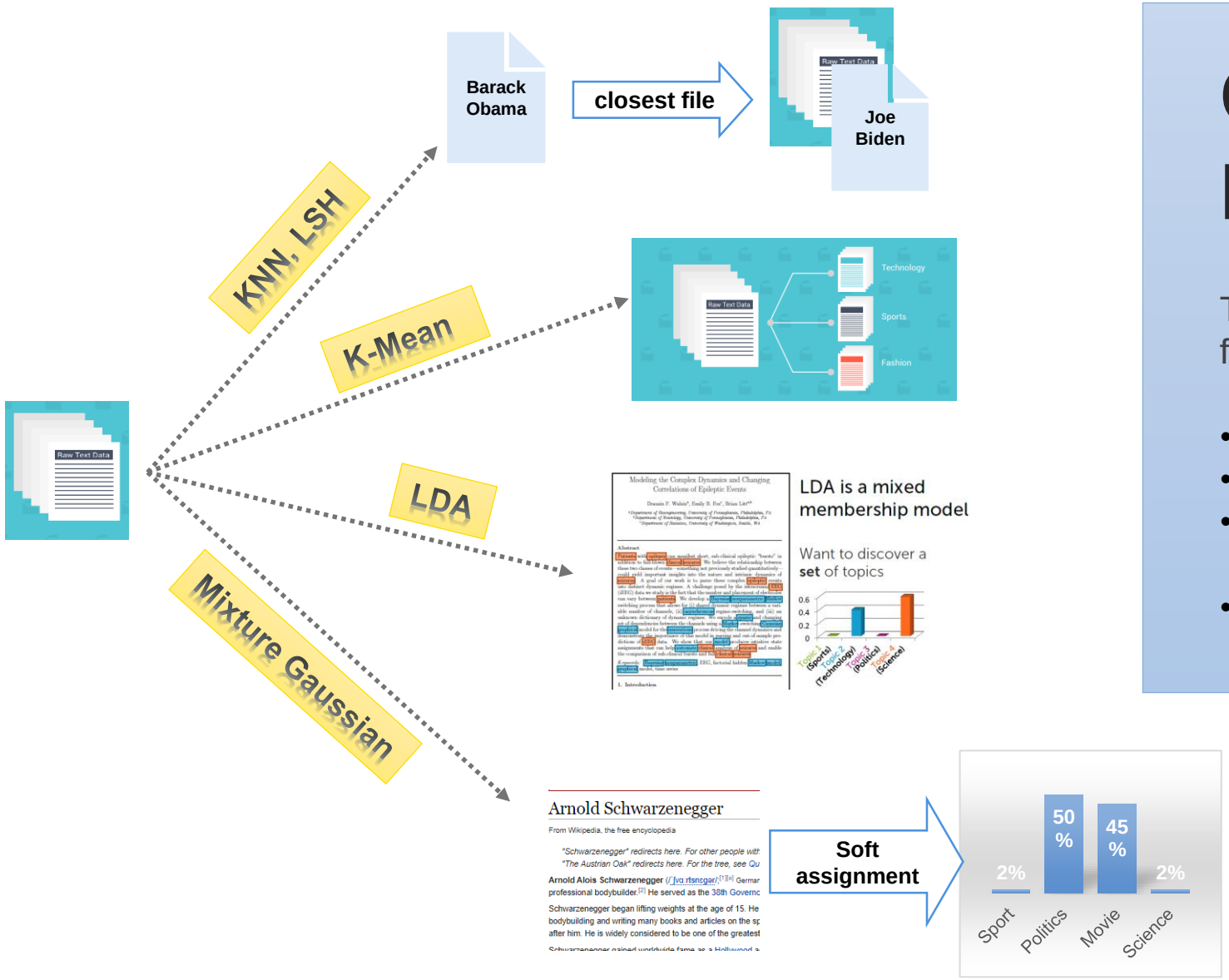
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	
id	membe	loan_a	funded	funded	term	int_rat	installr	grade	sub_gr	emp_t	emp_le	home_	annual	is_inc_	issue_d	loan_st	pymnt_	url	desc	purpos	title	z
1E+06	1E+06	5000	5000	4975	36 mor	10.65	162.9	B	B2		10+ yea	RENT	24000	Verifie	201112	Fully P	n	https:/	Borro	credit_	Compu	8
1E+06	1E+06	2500	2500	2500	60 mor	15.27	59.83	C	C4	Ryder	< 1 yea	RENT	30000	Source	201112	Charge	n	https:/	Borro	car	bike	3
1E+06	1E+06	2400	2400	2400	36 mor	15.96	84.33	C	C5		10+ yea	RENT	12252	Not Ve	201112	Fully P	n	https://www.l	small_l	real est	6	
1E+06	1E+06	10000	10000	10000	36 mor	13.49	339.3	C	C1	AIR RES	10+ yea	RENT	49200	Source	201112	Fully P	n	https:/	Borro	other	person	9
1E+06	1E+06	5000	5000	5000	36 mor	7.9	156.5	A	A4	Veolia	3 years	RENT	36000	Source	201112	Fully P	n	https://www.l	weddir	My we	8	
1E+06	1E+06	3000	3000	3000	36 mor	18.64	109.4	E	E1	MKC A	9 years	RENT	48000	Source	201112	Fully P	n	https:/	Borro	car	Car Do	9
1E+06	1E+06	5600	5600	5600	60 mor	21.28	152.4	F	F2		4 years	OWN	40000	Source	201112	Charge	n	https:/	Borro	small_l	Expand	9
1E+06	1E+06	5375	5375	5350	60 mor	12.69	121.5	B	B5	Starbu	< 1 yea	RENT	15000	Verifie	201112	Charge	n	https:/	Borro	other	Buildin	7
1E+06	1E+06	6500	6500	6500	60 mor	14.65	153.5	C	C3	Southw	5 years	OWN	72000	Not Ve	201112	Fully P	n	https:/	Borro	debt_c	High in	8
1E+06	1E+06	12000	12000	12000	36 mor	12.69	402.5	B	B5	UCLA	10+ yea	OWN	75000	Source	201112	Fully P	n	https://www.l	debt_c	Consol	9	

AdaBoost



Decision Tree





Cluster Wikipedia Documents

The task is to cluster similar documents together and find the topics.

- Project: Cluster Wikipedia Documents
- Data input: text
- Model Structure: **Tokenize, bag of word, TF-IDF, KNN, Kmean, LSH, LDA, Guassian distribution**
- Problem solved: nearest neighbor, probabilities, cluster

Recommendation System

The algorithm is to build up a movie recommendation system.

- Project: Movie recommendation system
- Model Structure: **Collaborative filtering, Content based recommendation**

Input

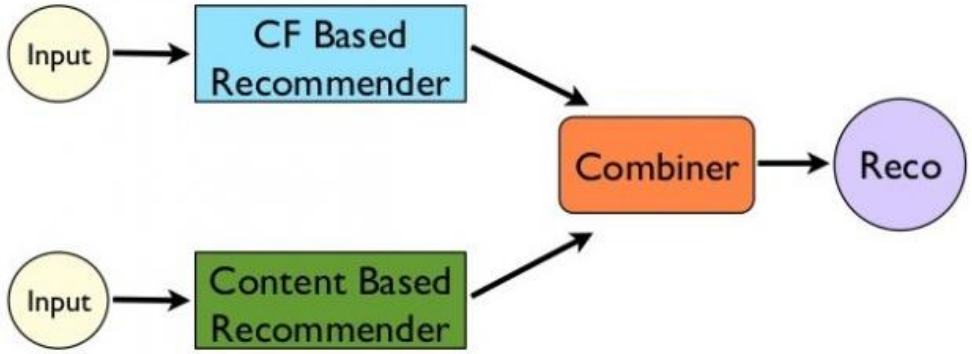
Content based

Movie	Alice (1)	Bob (2)	Carol (3)	Dave (4)	x_1 (romance)	x_2 (action)
Love at last	5	5	0	0	0.9	0
Romance forever	5	?	?	0	1.0	0.01
Cute puppies of love	?	4	0	?	0.99	
Nonstop car chases	0	0	5	4	0.1	
Swords vs. karate	0	0	5	?	0	

Collaborative filtering

Movie	Alice (1)	Bob (2)	Carol (3)	Dave (4)
Love at last	5	5	0	0
Romance forever	5	?	?	0
Cute puppies of love	?	4	0	?
Nonstop car chases	0	0	5	4
Swords vs. karate	0	0	5	?

Model



Output

Top recommendations for you:
Predicting rating 9.0 for movie Titanic (1997)
Predicting rating 8.9 for movie Star Wars (1977)
Predicting rating 8.8 for movie Shawshank Redemption, The (1994)
Predicting rating 8.5 for movie As Good As It Gets (1997)
Predicting rating 8.5 for movie Good Will Hunting (1997)
Predicting rating 8.5 for movie Usual Suspects, The (1995)
Predicting rating 8.5 for movie Schindler's List (1993)

Anomaly Detection

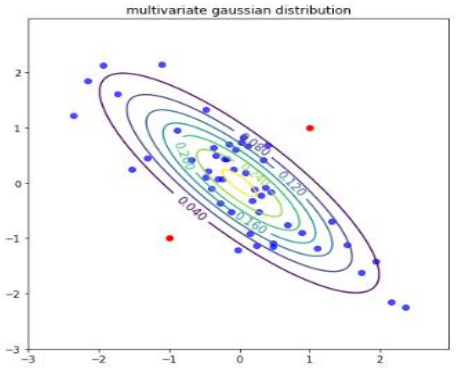
The project is to detect the abnormal data from all data.

- Project: Anomaly detection
- Statistic model: **Univariate Gaussian , Multivariate Gaussian**

Anomaly Detection

Statistic Model

All data



Anomaly detection algorithm

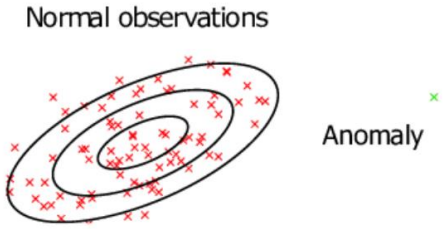
1. Choose n features x_i that you think might be indicative of anomalous examples.
2. Fit parameters $\mu_1, \ldots, \mu_n, \sigma_1^2, \ldots, \sigma_n^2$

$$\mu_j = \frac{1}{m}\sum_{i=1}^{m} x_j^{(i)} \qquad \sigma_j^2 = \frac{1}{m}\sum_{i=1}^{m}(x_j^{(i)} - \mu_j)^2$$

3. Given new example x , compute $p(x)$:

$$p(x) = \prod_{j=1}^{n} p(x_j; \mu_j, \sigma_j^2) = \prod_{j=1}^{n} \frac{1}{\sqrt{2\pi}\sigma_j} \exp\left(-\frac{(x_j - \mu_j)^2}{2\sigma_j^2}\right)$$

Anomaly if $p(x) < \varepsilon$



S U M M A R Y

Artificial Intelligence is Everywhere

We now live in the age of “big data,” an age in which we have the capacity to collect huge sums of information too cumbersome for a person to process. The application of artificial intelligence in this regard has already been quite fruitful in several industries such as technology, banking, marketing, and entertainment.
Deploying and exploring in the field will make a big difference in the world in the next twenty years, I cannot wait to be part of it.

Portfolio Summary

Autonomous Driving

Finance / Banking

Social Network

Entertainment

Marketing

Security

The fields are involved in my projects.

Thank You

Feel free to contact me if you have any questions

yanjialie@gmail.com