Data Analysis in SenseTime

Xinlin Hu Lifang Yu Xinyu Wang Ziwei Yuan Haohua Chen
Department of Mathematics
Hong Kong University of Science and Technology
Clear Water Bay, Hong Kong
{xhuax, lyual, xwangek, zyuanag, hchenbv}@connect.ust.hk

Abstract

Artificial Intelligence are applied in board fields and goes noticeable recent years. During the project, Skyfall team took the industrial leader, SenseTime, which focuses on computer vision and deep learning, as the research target. Fundamentally, industrial structure, competitors and are analysed with the supporting of data. Technically, Weibo and Twitter text from SenseTime and its competitors are scraped by using Python language. Afterwards, word clouds, word vectors, histograms and sentiment analysis verify SenseTime's innovation direction.

1 Technical report

1.1 Industrial structure and Competitor

1.1.1 Industrial structure

The industrial chain structure of artificial intelligence is divided into basic layer (computing infrastructure), technical layer (software algorithm and platform) and application layer (industry application and products).

The basic layer mainly includes computing hardware (AI chip), computing system technology (cloud computing, big data and 5G communication) and data (data collection, annotation and analysis).Take AI chips as an example. As the core hardware of the AI industry, some analysts believe that the market size of AI chips will reach 14.616 billion us dollars by 2020, accounting for about 12.18

At the technical level, we can understand from three dimensions: algorithm theory (machine learning algorithm, brain-like algorithm), development platform (basic open source framework, open technology platform) and application technology (computer vision, natural language understanding and human-computer interaction).

At the application level, it can be divided into industrial solutions (" AI+ ") and typical products (robots, smart speakers, smart cars, drones, etc.). Industrial solutions are mainly "AI+ traditional industries", which cover many vertical fields such as security, transportation, medical care, manufacturing, education, finance and home furnishing.

At present, the United States is still one of the core cradles of artificial intelligence, with excellent technology research and development institutions, theoretical disciplines and all kinds of laboratories. With favorable capital and policies, the development prospect of the AI industry is optimistic. China has become one of the global AI centers. However, data environment, talent shortage and the immaturity of intelligent hardware, especially microchips, CPU and other industries are arguably the biggest challenges facing the development of AI in China.

32nd Conference on Neural Information Processing Systems (NeurIPS 2018), Montréal, Canada.

1.1.2 Competitor

With the development of technology, great changes have taken place in Artificial Intelligence industry, especially in Computer Vision stream. There is no doubt that this critical high-technology may improve the social efficacy and lift quality of human being. More and more capitals, i.e venture capital and private equity, invested to the industry for revenues. For the project, hence the target financing entity, sensetime, is based on China, the other 5 competitors are selected to compared with it from different points of view. Respectively, they are Megvii, YITU, Cloudwalk, DeepBlue and Clobotics.

Table 1. Basic information for competitors

	Sensetime	Megvii	YITU	Cloudwalk	DeepBlue	Clobotics
Found date	Aug 2014	Oct 2011	Sep 2012	Mar 2015	Oct 2014	Nov 2016
Headquarter	Beijing, Shenzhen	Beijing	Shanghai	Chongqing	Shanghai	Shanghai
Valuation	$3 billion	$4.5 billion	$2.4 billion	¥22 billion	N/A	N/A
Business map	Defense, Finance, Retails, Automobile, Education	Defense, Finance, Retails, Automobile, Education, Logistic	Healthcare, Automobile	Defense, Finance,	Defense, Retails, Automobile, Education, Logistic	Retails, Energy

From the information shown above, it's not hard to know that Sensetime is not an eldest Computer Vision entity on the list, but it's counted for a high valuation, compared to the other competitors. In addition, many technical and basic fields are covered by Sensetime, which will be a great opportunity for it to go further in the emerging market.

Table 2. Business map for Computer Vision Entity

	Defence	Finance	Retails	Automobile	Education	Logistic	Healthcare	Energy
Sensetime	Face recognition	ID verification	Consumer analysis	Driver system	AI textbook			
Megvii				Driver recognition	Expression analysis	AI classification		
YITU		Face taking money					Care.AI platform	
Cloudwalk		Face paying						
DeepBlue			Non-cashier retail					
Clobotics			Monitor platform	Auto-driving	Handwrite recognition	AI logistic robots		Equipment management

Table 3. Different business fields of AI Entrepreneur entities

Field	Business Model	Suppliers	Competition situation
Internet	Software, 3D Optics	Megvii	Heavy competition stress from Internet entities
Defence	AI engine	Sensetime, Megvii,YITU,Cloudwalk,DeepBlue	High entry barrier
Finance	Face-matching	Sensetime, Megvii, YITU, Cloudwalk, DeepBlue	Low price sensitivity, high commercialization
Retails	Non-cashier, consumer analysis	Sensetime, Megvii, YITU, Cloudwalk	Limitation from soft and hardware
Automobile	Automatic driving	Sensetime, Megvii, DeepBlue	On initial stage, Share economy
Healthcare	Medical imaging system	YITU	Less competitor, high technical requirement

From Table 2 and Table 3, we can get the implication that all most all its competitors are arranging their business map. Sensetime and the other 5 industrial peers engages in defence by applying the technology of face recognition. But obviously, by comparing the variety of business, only Megvii and Clobotics might bring a huge competition threaten to Sensetime due to the diversified products. The lack of product in logistic, healthcare and energy is a weakness for Sensetime's operating.

1.2 Data mining

1.2.1 Publications analysis by wordcloud

Due to figure out the trend of research focuses of SenseTime, we collected their information of patents, software copyrights and academic publications from official websites such as CVPR and TianYanCha. After browsing those website, we found that SenseTime have over 130 software copyrights, and over 250 patents. To visually illustrate the trend of research direction, we used R to draw the word cloud of those information, finding the top-frequent words in the description of those patents, copyrights and publications.

Firstly, we formed three word clouds to represent the change of its patents from 2017 to 2019, respectively. In 2017, the key words of patents are related to interface of devices and face recognition, which means that they paid attention to both outer design and algorithms. In 2018, company attached more importance with face recognition and relevant network applications which could be reflected by the occurrence of words like 'shi pin', 'chan pin', 'yong hu'. In 2019, applications of their technology are gradually given the top priority.

(a) patents in 2017 (b) patents in 2018 (c) patents in 2019

Figure 1: Word cloud of patents from Sensetime

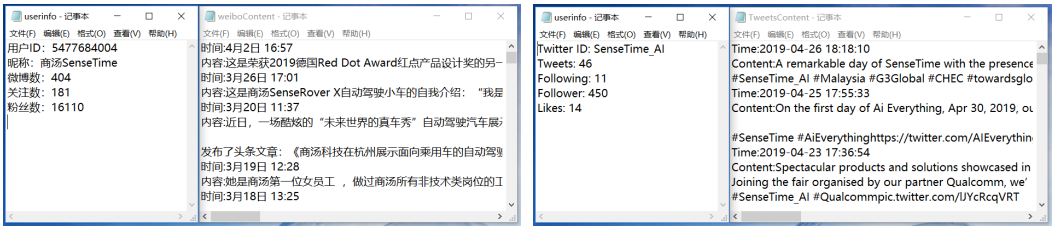
Secondly, we also used the summary of academic publications we found on CVPR and the information of software copyrights to form word cloud pictures. From those pictures, the most frequent words are 'network', 'feature', 'face', 'neural', 'video', 'human face', 'AI', which are consistent with what we learned from its patents.

(a) Publications in CVPR 2018 (b) Software

Figure 2: Word cloud of publications from Sensetime

1.2.2 Web scrapying

We use the same module called 'selenium' in python to do web scrapying task in Weibo and Twitter to get all the weibo contents and tweets from Sensetime. Firstly, we download 'ChromeDriver' and put it in the directory as 'C:/Program Files (x86)/Google/Chrome/Application'. Then we program with 'selenium' module to simulate how people use computer to copy weibo contents or tweets to laptop. The codes have been shared in Github[1][2].

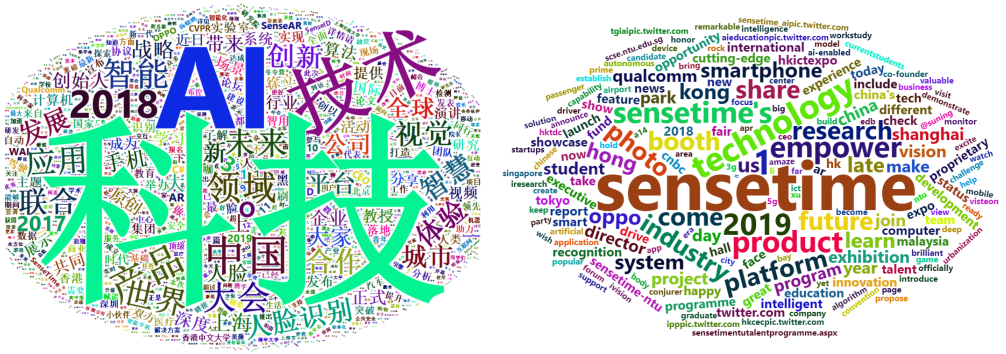


(a) Weibo's profile and contents (b) Twitter's profile and tweets

Figure 3: Word cloud of publications from Sensetime

We successfully get the profile of Sensetime in weibo and twitter, we also get all the weibo contents and tweets in our laptop, in order to analyze the company developments in the later study.

Word cloud After getting weibo contents and tweets, we also draw out the word cloud.

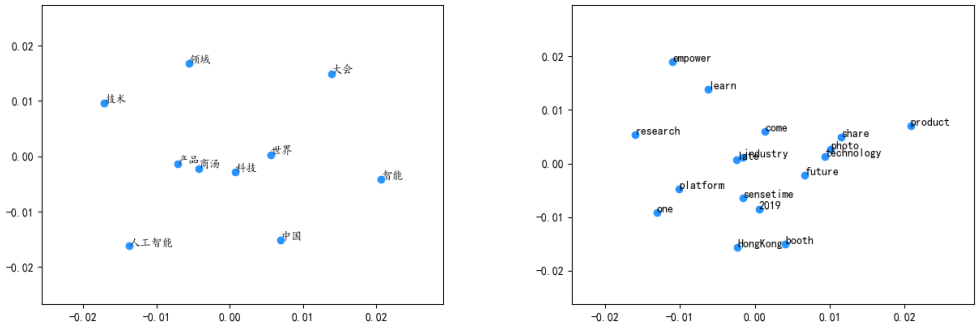


(a) Weibo contents (b) Tweets

Figure 4: Word cloud of weibo contents and tweets

Word vector We directly use the word vectors from Github trained by Shen Li, et al. (2018), and we use PCA to reduce the dimension of word vectors by calling 'from gensim.models import Word2Vec' in python to draw out the figures.

[1]WeiboSpyder: https://github.com/thompsonhu/weiboSpyder
[2]TwitterSpyder: https://github.com/thompsonhu/twitterSpyder



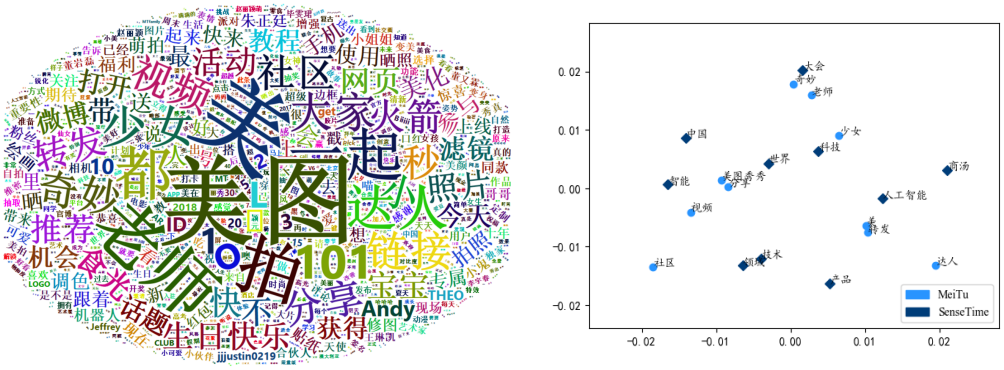
(a) Weibo contents (b) Tweets

Figure 5: Word vector of weibo contents and tweets

1.2.3 Analysis of competitor's social account information

Nowadays, posting status on social media and interacting with its followers is regarded as one of the most efficient methods to realize brand promotion. Due to this situation, we could explore more potential information from the content posted by their official social accounts, both its brand image and its recent focus. Meanwhile, it is not enough to analyze only the information of SenseTime, its competitors and partners are crucial as well. So, we used Creeper code on Python to obtain all the statuses posted by Megvii (competitor) and Meitu (partner). Then we used those contents to form word cloud and word vector model to compare the the differences of those companies.

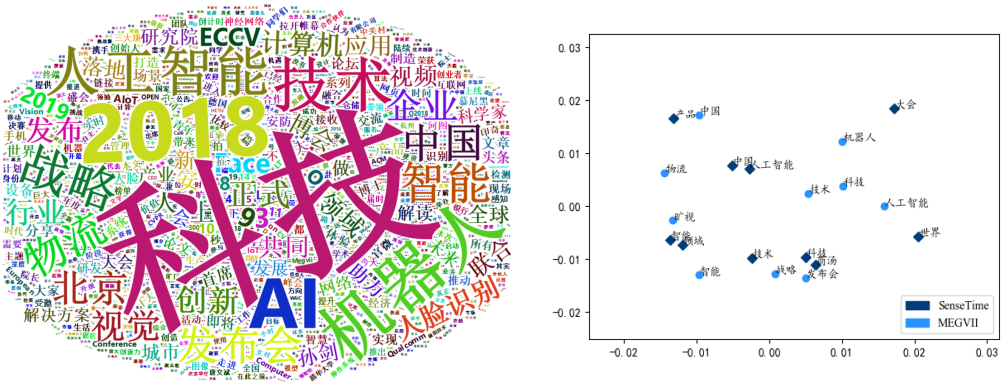
For Meitu, its weibo account is significantly active than other two companies, with more than 30,000 posts. The key words are closely related to video, online community, photograph and young girls. From the comparison of word vectors between SenseTime and Meitu, there is a close relation among those words, which means that our objective company well caters to the demand of its clients.



(a) Word cloud from Meitu's weibo (b) Meitu vs Sensetime

Figure 6: Data analysis on Sensetime with its partner

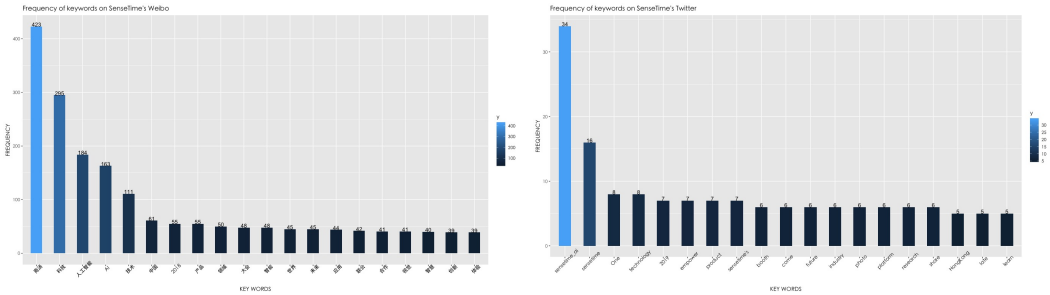
For Megvii, we could see from its word cloud and word vector chart that technology, AI, visual and robot are its key points, this shares some common parts with SenseTime, but they do exist some differences. While SenseTime pays more attention to monitoring equipments, its competitor keeps a watchful eye on robots and logistics technology.



(a) Word cloud from Megvii's weibo (b) Megvii vs Sensetime

Figure 7: Data analysis on Sensetime with its competitor

To show the hot topic implied by SenseTime's social account, we also counted the top-frequent words contained in its Weibo and Twitter. It is not surprising that 'SenseTime', 'ke ji', 'ren gong zhi neng', 'ji shu' occupy the top places of this list and left other words far behind in Weibo, and words like 'SenseTime', 'technology', 'product', 'research' appears on the list of Twitter. The two histograms shows a significant similarity. Then, we can observe that this company also mentions visual sense, future, cooperation and creativity very often. All those reflects that they are always trying to catch up with the most updated trend in AI field.



(a) Frequency of words from Weibo (b) Frequency of words from Twitter

Figure 8: Frequency of words

1.2.4 Sentiment Analysis

We do a simple demo of sentiment analysis on each weibo of SenseTime. Here we use python library SnowNLP, which is a package that can handle Chinese content conveniently. The model that has been trained in this package is based on the product's comment data which is suitable for the weibo content of SenseTime. Then we traverse every weibo content and call SnowNLP function s.sentimens() for sentiment analysis. Higher score means the content more optimistic . Here shows a pie chart of positive and negative emotions (see Fig. 9).

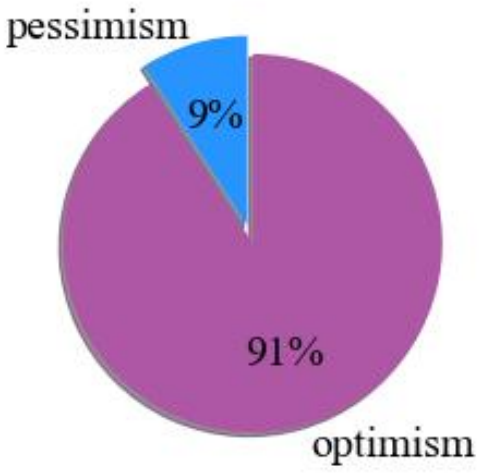


Figure 9: Pie chart of ratio for positive and negative emotions

2 Essay summary

2.1 AI essay - You Only Look Once: Unified, Real-Time Object Detection

2.1.1 Introduction

Object detection is one of the most important field in computer vision, whose goal is to 'teach' computer to figure out the objects in picture or any frame in video and recognize them. This is an interesting and challenging task. YOLO (You Only Look Once) is an object detection system targeted for real time processing which is introduced by Joseph Redmon, Santosh Divvala and etc. in 2016. The whole system contains the training part and test part.

2.1.2 Method

Training part The whole YOLO system contains 24 convolutional layers, 2 fully connected layers and 4 maxpool layers. In this architecture, they use leaky relu function as activation function. The output of YOLO system is a $7 \times 7 \times 30$ tensor which can be considered as 49 (7×7) grid cells and each grid cell contains 30 channels. These 30 channels will store the information of 2 bounding boxes (2×5, where 1 channel for box confidence and 4 channels for coordinate information) and 20 class probabilities (as Figure1 showed). No matter how many bounding boxes we create, we only predict one set of class probabilities for one time. Actually in practice, you can store those information in different way as you want.

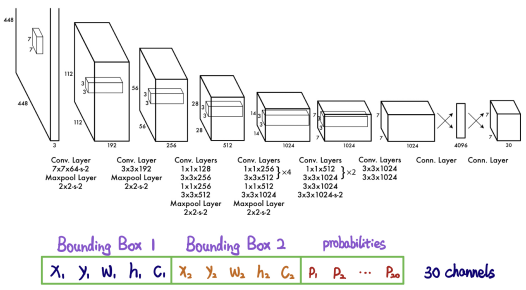


Figure 10: YOLO achitecture

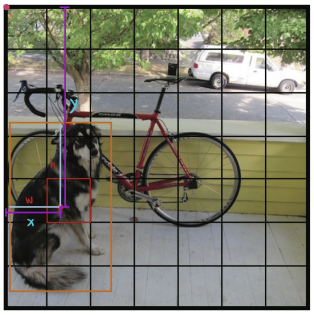


Figure 11: Coodinate information

As Figure10 showed above, YOLO used four variable ('x', 'y', 'w' and 'h') to locate the bounding box. They represents the coordinate of center point in bounding box and the size of bounding box, respectively. Every box will be given a confidence value, and the confidence is defined as:

$$\text{confidence} = \Pr(\text{Object}) \times \text{IOU}_{\text{pred}}^{\text{truth}}$$

Loss function The loss function can be devided into three parts: the coordinate error, IOU error and classification error. All the errors are measured by sum of square. The multi-part loss function is defined as:

$$\begin{aligned}
&\lambda_{\textbf{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^{B} \mathbb{1}_{ij}^{\text{obj}} \left[(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 \right] \\
&+ \lambda_{\textbf{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^{B} \mathbb{1}_{ij}^{\text{obj}} \left[\left(\sqrt{w_i} - \sqrt{\hat{w}_i} \right)^2 + \left(\sqrt{h_i} - \sqrt{\hat{h}_i} \right)^2 \right] \\
&+ \sum_{i=0}^{S^2} \sum_{j=0}^{B} \mathbb{1}_{ij}^{\text{obj}} \left(C_i - \hat{C}_i \right)^2 \\
&+ \lambda_{\text{noobj}} \sum_{i=0}^{S^2} \sum_{j=0}^{B} \mathbb{1}_{ij}^{\text{noobj}} \left(C_i - \hat{C}_i \right)^2 \\
&+ \sum_{i=0}^{S^2} \mathbb{1}_{i}^{\text{obj}} \sum_{c \in \text{classes}} \left(p_i(c) - \hat{p}_i(c) \right)^2
\end{aligned}$$

where $\mathbb{1}_i^{\text{obj}}$ denotes if object appears in cell i and $\mathbb{1}_{\text{ij}}^{\text{obj}}$ denotes that the jth bounding box predictor in cell i is "responsible" for that prediction.

The square root in second term can reduce the error when we compare the larger object with the smaller one. Also, paper choose coord = 5 and noobj = 0.5 in order to reduce the gradient from those grid cells do contain object.

Detection part In detection part, we mainly use the conditional probability formula:

$$\Pr(\text{Class}_\text{i}|\Pr(\text{Object})) * \Pr(\text{Object}) * \text{IOU}_{\text{pred}}^{\text{truth}} = \Pr(\text{Class}_\text{i}) * \text{IOU}_{\text{pred}}^{\text{truth}}$$

We can obtain the final detection box by setting threshold value to change low score value into zero and using NMS (Non-Max Suppression) to remove redundant bounding boxes.

Figure12 showed how to deal with several bounding boxes by NMS.

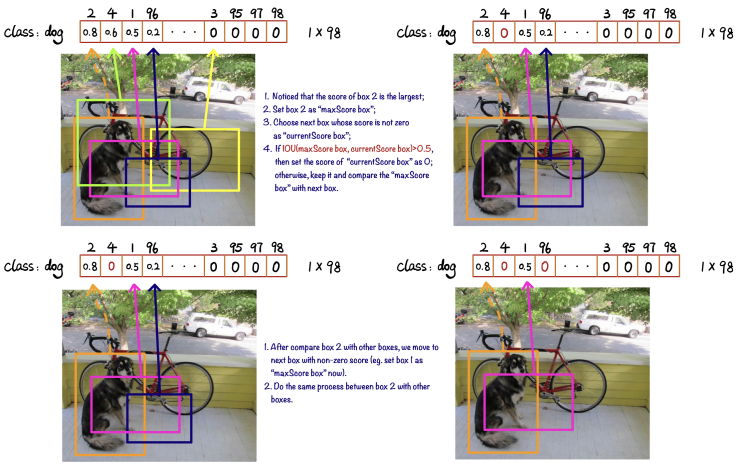


Figure 12: Removing redundant BBoxes by NMS

Detection result In detection part, we set the score threshold as 0.15 and IOU threshold as 0.6. We showed how detected result changed when we trained with more epochs.

Figure 13: Detection result

2.1.3 Conclusion

We noticed that YOLO version 1 is hard to detect the object when they overlapped together and this is the drawback of YOLO version 1, but YOLO later version solved this problem.

2.2 Finance essay- Artificial intelligence and machine learning in financial services Market developments and financial stability implications

2.2.1 Applications

From this essay, we find that AI is being adopted for a number of purposes across the financial system. For example, the back-office applications of AI and machine learning could improve risk management, fraud detection, and compliance with regulatory requirements, potentially at lower cost. In portfolio management, the more efficient processing of information from AI and machine learning applications could help to boost the efficiency and resilience of financial markets–reducing price misalignments earlier and reducing crowded trades. Nonetheless, use of AI and machine learning risks creating 'black boxes' in decision-making that could create complicated issues, especially during tail events. In particular, it may be difficult for human users at financial institutions – and for regulators – to grasp how decisions, such as those for trading and investment, have been formulated. Finally, with use cases by regulators and supervisors, there is potential to increase supervisory effectiveness and perform better systemic risk analysis in financial markets.

As the underlying technologies develop further, there is potential for more widespread use, beyond the use cases discussed in this essay. It will be important to continue monitoring these innovations and to update this assessment in the future.

2.2.2 Privacy concerns

The issues around data privacy relate to the ability to access the data being processed by AI and machine learning tools. While big data are widely used to generate profits, they can only do so with technology that converts the data into relevant services.1 In financial services, AI and machine learning applications usually depend on access to, and use of, large amounts of data in a 'life cycle' that includes data collection, data compilation and consolidation, data mining and analytics. AI and machine learning present a range of legal issues relating to privacy and data protection, consumer protection, anti-discrimination and liability issues, and cross-border issues.

3 Suggestion for further study

To team members:

- Be proactive to refer more relevant materials.
- Try more different methods to scrap source and clean data, i.e. not limited in Python but includes R, C++, Java.

- Set a timetable and timeline after finishing group meeting and before next steps.
- Define every member's role in the project according to their own skills.
- Share all discovery or experience during group meeting, which let them teach each other.

To tutors:

- Be better to increase the frequency of consultation on campus.
- Improve the connection and relationship between theoretical part and project part in class.
- Try to invite more professional and industrial experts to give student precious career talks.

Acknowledgments

- HU Xinlin - Scrape Weibo and Twitter by Python, Summurize report, Edit video, Record screen
- WANG Xinyu - Collect data, Prepare slides, Senstiment Analysis, Visualize word vectors, Record screen
- YU Lifng - Prepare slides, Present project, Analyze competitors, Arrange matters
- YUAN Ziwei - Collect data, Plot word cloud and vecetors, Present project
- CHEN Haohua - Edit video, Analyze indusrty, Summurize report, Present project

References

[1] Shen Li, Zhe Zhao, Renfen Hu, Wensi Li, Tao Liu, Xiaoyong Du, (2018) Analogical Reasoning on Chinese Morphological and Semantic Relations, *ACL.*

[2] Bower, J.M. & Beeman, D. (1995) *The Book of GENESIS: Exploring Realistic Neural Models with the GEneral NEural SImulation System.* New York: TELOS/Springer–Verlag.

[3] Hasselmo, M.E., Schnell, E. & Barkai, E. (1995) Dynamics of learning and recall at excitatory recurrent synapses and cholinergic modulation in rat hippocampal region CA3. *Journal of Neuroscience* **15**(7):5249-5262.