Scrapnet: A Blockchain and Machine Learning based System to Facilitate Scrapping of Cars

Rehana Zenab
Dept. of Computer Science and EngineeringMar Athanasius College of Engineering Kothamangalam, Kerala

Sreerenjini R Chandran
Dept. of Computer Science and EngineeringMar Athanasius College of Engineering Kothamangalam, Kerala

Vidhya
Dept. of Computer Science and EngineeringMar Athanasius College of Engineering Kothamangalam, Kerala

Prof. Neethu Subash
Dept. of Computer Science and EngineeringMar Athanasius College of Engineering Kothamangalam, Kerala

***Abstract*—Scrapping of cars in India is not an organized process like the sale of used cars. Transactions related to vehicles mainly involve buying and selling. However, tracking these transactions can be a challenging task. Even though there exist some solutions using centralized systems, they may have problems with transparency, trust, and access control. Also, scrapdealers might deceive naive customers by fixing unfair pricesfor the cars given for scrapping. Therefore, in this paper, we provide an integrated blockchain and machine learning-based solution for automating the transactions related to the scrapping of cars. The major parties involved in the transactions are the Regional Transport Officer (RTO), car owners, and scrap dealers,and their communication is facilitated with the help of smart contracts. Along with this, damage detection followed by price prediction of the car to be scrapped is carried out which helps thecar owner realize the actual worth of his/ her car without being deceived. We propose Extreme gradient boosting (Xgboost) model for the same. The obtained results show that Xgboost regressor achieves high-performance gains, having an accuracy of 89.96 percent. Finally, a scrapping certificate is generated and issuedto the owners of the scrapped cars, with which they can avail discount while buying a new car thereby attracting more people to get their old and pollution-causing cars scrapped.**

***Keywords*—Scrapping, Blockchain, Machine learning, Smart Contracts**

I. Introduction

A car starts degrading and releases toxic emissions either due to an accident or aging, which in turn contaminates the environment. Such a car is considered a scrap car according to the Scrapping policy. In addition, as the car ages, it will need extensive maintenance and repairs, which will increase the cost of ownership. In India, a new vehicle scrappage policy[1] was recently announced, under which commercial vehicles over fifteen years old and private vehicles over twenty years old are required to be scrapped if it fails fitness and emission tests. Scrapping ensures that the vehicle is dismantled into pieces and recycled. It also eliminates the possibility of the vehicle being used for unwanted or illegal purposes.

This paper concentrates on the scrapping of cars. There are some protocols [1] associated with car scrapping. According to the norms, the car owner should approach the authorized Regional Transport Officer (RTO) and submit a letter expressing the intent to scrap the car. Once an agreement has been reached, the scrap dealer will dismantle the car partsand segregate them into iron, plastic, rubber, etc. During the scrapping process, the chassis number of the car is cut out before the scrap is sent for recycling. This chassis number has to be submitted to RTO as part of the deregistration process. After getting the vehicle scrapped, the car owner is supposed to submit the registration certificate along with the chassis number that is cut out of the car while scrapping. The car owner will also have to submit an affidavit mentioning that thecar is not under any loans, insurance claims or pending court cases. It should also mention that the car is not involved in anykind of criminal activity such as robbery. Once the documents are verified by the RTO, the car will be deregistered.

Even though such rules exist, many people are unaware of these legal procedures[1] to be followed for getting their vehicles scrapped. This traditional method is complex and centralized. It requires a huge amount of time and manual work which is one of the biggest drawbacks in today’s world of digital technologies. Therefore, in this paper, we developa secure and automated system framework [2] to implement scrapping of cars to reduce human interaction. To the best of our knowledge, there is no previous work using blockchain and machine learning technologies for this application.

The proposed architecture makes use of the strengths of blockchain [3] to ensure the integrity of cars provided for scrapping and to provide participants with a tamper-proof infrastructure for conducting various transactions [4]. In our system, a person who wishes to get his or her old car scrapped can do so, following the legal and standard procedure for scrapping. The car owner can interact with an authorized scrap dealer and can ensure that he or she receives the correct amount for the scrapped car. The system also attracts more people to give away their old cars for scrapping by providinga scrapping certificate to car owners with which they can availdiscount on buying a new car, thereby reducing pollution and other environmental impacts caused by old cars.

In addition, we propose to employ machine learning methods to build the damage detection and price prediction modules. There exist cases where certain scrap dealers deceive innocent car owners by fixing a lesser price for the cars given for scrapping. The car owners will fall for it believing the dealer's expertise. In order to avoid this situation, the price is automatically predicted in our system. Detecting the amount of outer damage a car posses, is an important step that precedes price prediction. The damage rate is found out and along with it, many other features of cars and their associated cost are fed as input to the price prediction models which predicts the cost of the car as output.

The rest of the paper is organized as follows: Section II provides a literature review about different fields related to scrapping and recycling of cars. In Section III, we describe system design of the proposed network based on blockchain and machine learning technologies. Machine learning methodologies are explained in detail in Section IV. The obtained results are presented in Section V. The conclusions and future work are depicted at the end in Section VI.

II. Related Works

There have been numerous research works on vehicle dismantling and scrappage rate detection over the years. An example of the same is Automotive Recycling Information Management Based on the Internet of Things and RFID Technology. The Authors Tongzhu ZHANG, Xueping WANG, Jiangwei CHUI, Xianghai LIU, Pengfei CUll have introduced in [5] a system that promotes automotive recycling in China. Taking into account, the transportation safety and environmental protection, the Chinese government banned the reuse of end-of-life vehicles. When a vehicle reaches its ending stage, the owner chooses to resell or reuse the vehicle rather than handing over the vehicle to dismantlers. This leads to severe traffic safety issues and environmental pollution. The supply system of the automotive lifecycle consists of three stages, the beginning of life stage, middle of life stage, and end of life stage. The recycling information is disseminated to the producers to improve the product design and the consumers can seek the required components on the information server. This paper portraits the general model for managing recycling information and information flow at each phase of a vehicle lifecycle, to promote the development of the automotive products recycling industry in China. Authors Jie Li1, Michael Barwood1, and Shahin Rahimifard1 in [6] introduced a novel approach for recycling electric vehicles. Electric vehicles have been in high demand as people are indulging in personal transportation that results in air pollution. The current procedure of end-of-life vehicle recycling is followed by the stages,depollution, dismantling, shredding, post-shredder separation,and incineration. In order to reduce material contaminationand escalate value recovery from complex materials, two techniques are proposed. The first technique deals with the dismantling of valuable components and segregating themfrom the main vehicle. These segregated parts then undergo a similar recycling procedure which includes fragmentation and post-shredder separation. The second techniques consistof further disassembling the components and extracting the subparts. This paper also introduces a novel approach for recycling electric vehicles using semi-automated robotic disassembly. The proposed system supports the identification of essential components and segregating them in order to improve the quality of recycled materials and improve the economic potential and lower the impact of such vehicle parts on the Environment.

Authors Mohammad Z.Masoud, Yousef Jaradat, Ismael Jannoud, and Dema Zaidan in [7] aimed at Buying and selling used vehicles which is one of the popular businesses acrossthe world. While purchasing used vehicles, customers are concerned about the car's usage, intent for sale, condition, etc. In order to overcome this, a blockchain-based framework has been introduced to report the history of used vehicles. It is a peer-to-peer distributed application. In the proposed system different users like car owners, common people, car repairing organizations can join or register. The system is built upon various smart contracts for registering new users and updating the blocks. The proposed system allows car owners to grant permission to view the car history report.

Authors Samiksha Marne, Shweta Churi, Delisa Correia, Joanne Gomes in [8] have highlighted the use of a Recurrent Neural Network (RNN) that used a Long Short Term Memory (LSTM) algorithm for predicting the value of cryptocurrency. The system comprises two main phases. The first one being visualizing and analyzing the bitcoin dataset and the second phase is the implementation of the RNN model with the help of the LSTM algorithm. The results were estimated by plotting graphs followed by root mean square error. First, the datasets are collected using online platforms like Kaggle. Secondly, the deep learning environment is set up. RNN is a type of deep learning method which is most prominently used for extracting patterns of temporal sequences which is effective in predicting bitcoin price. Finally, the data is fed to the LSTM regression model.

The authors Rishabh Ranka, Niranjan Sharma, Naman Talati, Nitika Rai in [9] aimed to implement a system to overcome the issues of direct taxes. The smart contracts are deployed to estimate taxes and to adjust the amount that has to be refunded. There are three entities in the system, Buyer, Seller, and the blockchain smart contract[3]. The buyer has to register to the network, after logging into the system, the buyer must add the details regarding purchasing the buyer's data is validated using the smart contract. The second entity seller is supposed to register onto the system and has to enter the details of the user to whom the payment has to be made. The smart contract [3] is responsible for verifying the data and allowing the users onto the network and permit them to perform transactions.

III. Proposed System Overview

The architecture of Scrapnet is given in figure 1. Our system Scrapnet helps people acquaint themselves with the impor-

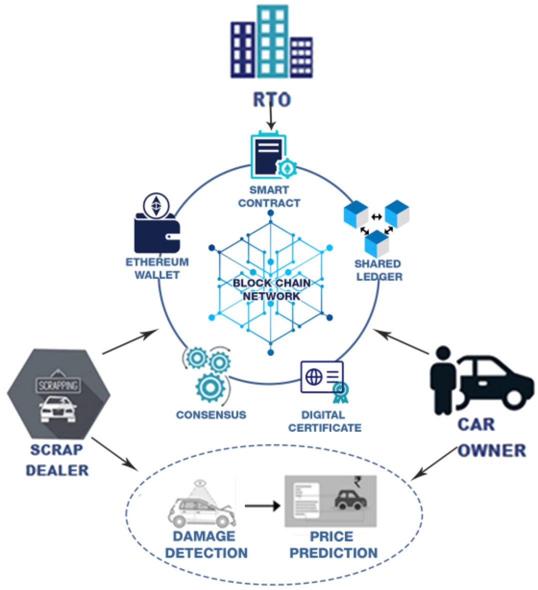

Fig. 1: Proposed system

tance of scrapping and encourages an easy scrapping experience. The proposed system comprehends a private Ethereum blockchain and machine learning technologies. The blockchain technology helps in acquiring a decentralized system. It leads to quick and more translucent settlements, as the ledger is automatically updated and can be accessed by each member of the network [3]. The system is incorporated with three entities, the scrap dealer, car owner, and the RTO. The scrap dealer, car owner, and RTO cooperate with each other using a proof of work consensus algorithm. Firstly, the scrap dealer registers with his details and wait for approval from the RTO, who in turn authenticates the scrap dealer's details and allows the scrap dealer into the network. Secondly, the car owner has to register with his details and wait for approval by the scrap dealer as well as verification by RTO. On verifying boththe scrap dealer and the car owner, RTO will approve the scrapping request.

The next step is detecting damage of the car and predicting the price [10] conforming to the damage detected and other features of the car. After scrapping the car, the scrap dealer has to put forward all information about the scrapped car to the RTO. The RTO will verify the details and issue a certificate to the car owner stating that the vehicle is deregistered. The certificate is generated using blockchain, each certificate has a unique hash value, based on which the authenticity of transactions can be monitored. The system revolves around four smart contracts [4] for car owner registration, scrap dealer registration, RTO verification and certificate generation. The smart contracts are deployed in a local environment. After compiling the smart contract, the resulting Application Binary Interface (ABI) must be integrated into the desired platform. The web or mobile interface that communicates with this application is developed using React. It is a JavaScript library

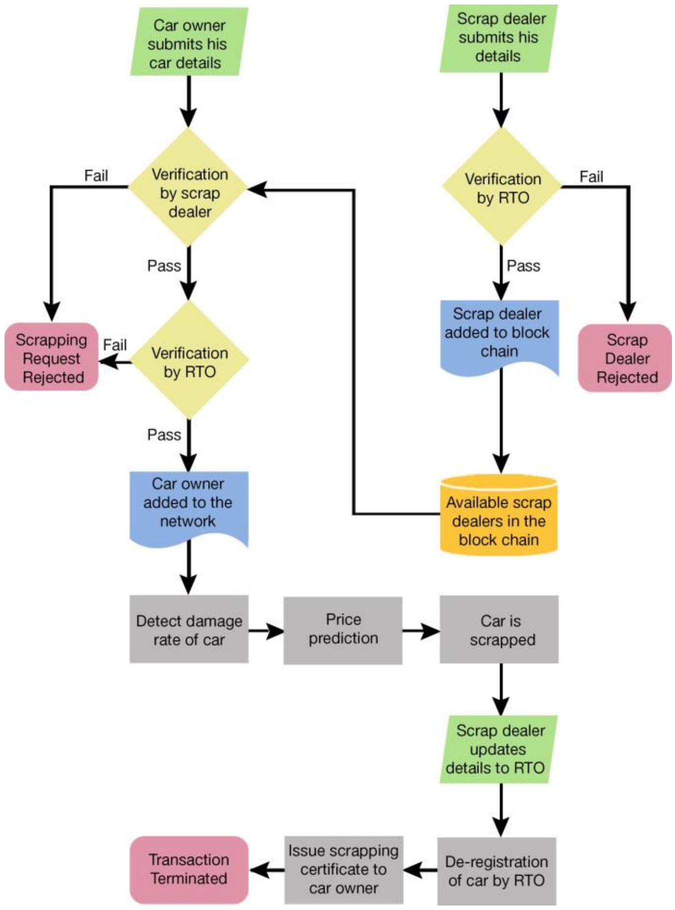

Fig. 2: Steps in Scrapping of car

designed for building swift and immersive user interfaces for web and mobile applications. This open-source front-end library is liable for the application's view layer only. The detailed workflow of our system is presented in figure 2.

IV. MACHINE LEARNING BASED DAMAGE DETECTION AND PRICE PREDICTION

The proposed methodology consists of collecting, cleaning and preprocessing the data, and implementing the models. Figure 3 summarizes the overview of this process.

A. *Data Collection*

The datasets used are collected from Kaggle. For damage detection, the dataset of car images are divided into 11 folders (0-10), 0 representing images of cars with no damage and 10 representing images of completely damaged cars. The dataset for price prediction involves the different attributes of a car such as id, region, year, manufacturer, model, paint color, image, description, type, price.

B. *Data Cleaning*

The first step in data cleaning is to remove irrelevant features like 'region', 'description' from our dataset. Next step is to check for missing values and fill them with appropriate values by an appropriate method.

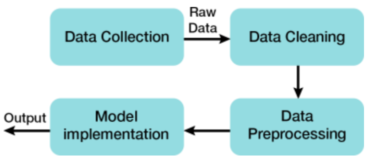

Fig. 3: Proposed machine learning module pipeline

ExtraTreesRegressor estimator is better for the imputation method to deal with missing values. In order to remove the outliers from the data, InterQuartile Range (IQR) method is used.

C. Data Processing

The dataset for damage detection is augmented with affine transformed images which improve the generalization performance of the classifier. Hence, we enlarged the dataset by appending it with random rotations and horizontal flip transformations. ImageDataGenerator class of Keras is used for the same. In case of price prediction, the dataset is classified into numerical and categorical columns. These categorical variables have to be converted into numerical variables in order to implement model. The sklearn library LabelEncoder is used for the same. The dataset is then normalized using sklearn library MinMaxScaler.

D. Model Implementation

1) Convolutional Neural Network (CNN) : The model of damage detection is employed using the CNN [12] algorithm. The figure 4 shows the neural network with different convolution and fully-connected layers. In convolution layers, we are extracting the features [12] from the images. In theselayers, the association between the pixels of an image is being preserved. Strides refer to the amount of pixels shift over the input matrix. In this model, the stride value is 2 which suggests we are moving the filters to 2 pixels at a time. Rectified Linear Unit is used activation function in deep learning models. It will output the input directly if it's positive, otherwise, it'll output zero. It is written as

$$f(x) = max(0, x) \quad (1)$$

where f is output as a function and x is an input. Pooling Layer- It is used when the input image is just overlarge. It reduces the amount of parameters of an oversized image. We have used Max Pooling in our model, which takes the most important element from the rectified feature map. The layer which we called a fully-connected layer is like the way the neurons are arranged in a very traditional neural network.We flattened our matrix-vector and feed it into a completely connected layer kind of a neural network. Finally, Softmax is used as an activation function to classify the outputs into 11 categories. It is defined as:

$$S(x_i) = \frac{\exp(x_i)}{\sum_j \exp(x_j)} \quad (2)$$

where x represents values from the neurons of the output layer, exp acts as the non-linear function, which is applied on x_i. The result of this function will never be zero, but it has a value near to zero. $\sum_j \exp(x_j)$ is a normalization term. All the values of exponential function is divided by the sum of exponential values and then convert them into probabilities.

2) XgBoost Regressor : XGBoost is a powerful implementation of gradient boosted decision trees [13]. It is particularly designed to optimize memory usage and utilize hardware computing power. The main goal [13] of boosting is to sequentially construct sub-trees from an indigenous tree such that every subsequent tree decreases the errors of the previous one. In this manner, the new sub-trees will update the previous residuals to scale back the value function error. Consider the dataset D

$$D = \{(x_i, y_i)\ where\ x_i \in R^m, y_i \in R\} \quad (3)$$

where m is the total number of features in x_i and y_i is the ground-truth of the sample i. The term n represents the number of samples in such a way that $|D| = n$ where $|.|$ refersto cardinality of a dataset which is the number of rows in the dataset.

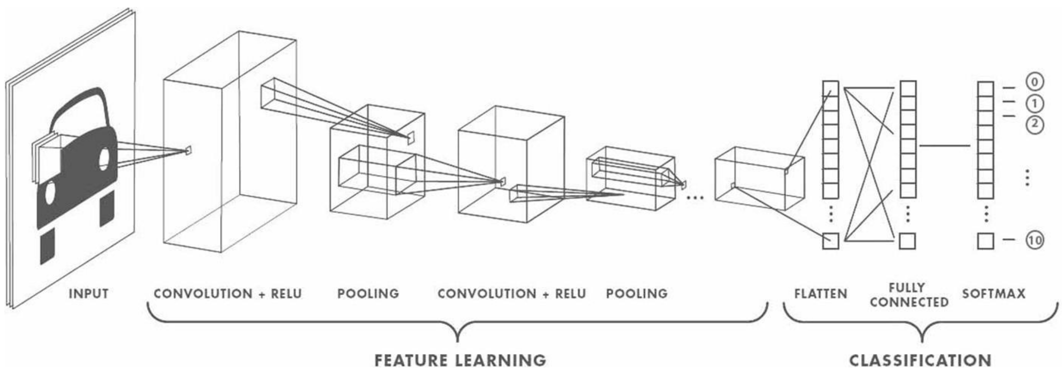

Fig. 4: A Convolutional Neural Network sequence to classify damaged cars

The predicted value of the entry i, $\hat{y_i}$, as follows:

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i), f_k \in F \quad (4)$$

where f_k denotes an independent tree in F, regression trees' space, and $f_k(x_i)$ indicates the predicted score provided by thei-th sample and k-th tree. The XGBoost cost function L can be expressed as follows:

$$\mathcal{L} = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k) \quad (5)$$

The training loss function l $(y_i, \hat{y_i})$ calculates the difference between prediction $\hat{y_i}$ and actual value y_i while $\Omega()$ is the regularization factor of the cost function; whose goal is to reduce the model risk in overfitting to the data. The factor $\Omega()$ can be represented as follows:

$$\Omega(f_k) = \Upsilon T + \frac{1}{2}\lambda \parallel w \parallel^2 \quad (6)$$

where Υ and λ are the two parameters[12] of regularization, and T and w denote the numbers of leaves and the scores of each leaf, respectively. By minimizing the L which is objective function, the regression tree model functions f_k can be studied. Learning rate is a parameter that determines what proportion we are adjusting the weights in our network which concernsthe loss gradient. The smaller the value, the slower we movealong the downward slope. The $n_e stimators$ refers to the num-ber of trees which will be created before taking the utmost vot-ing or averages of total predictions. In our model,200decision trees are built of 24 maximum depth and the learningrate is 0.4 with which the model learns. The price predictionmodel is trained for learning various features of cars, detected dam-age rate and their associated cost. The prediction is based up-on the extensive training dataset and hence the output of the model is the estimated value of the car.

V. Results

The damage rate was detected using CNN model with an accuracy of 45 percent. Also, it was found that the features of the car like age, odometer and damage rate have a major influence on the price predicted. The performance of Xgboost regressor model for price prediction was compared with other machine learning models and accuracy plot between the mod-els is shown in the figure 5. It can be seen that XgBoost regressor model provided better performance when compared to others, with an accuracy of 89.96 percent.

VI. Conclusion and Future Work

Creating an immutable ledger is one of the predominant features of blockchain, which helps to achieve a system that is tamper-proof and trustworthy. In this paper, we called attention to the importance of Car Scrapping thereby providing a hassle-free platform to interact with the scrap dealer and RTO. We proposed a solution to prevent manual

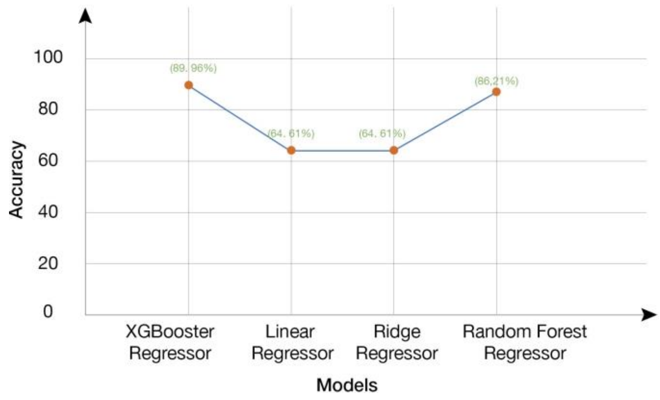

Fig. 5: Accuracy plot

work by generating certificates using blockchain, and detecting damage, and predicting the price using Machinelearning. We clarified the use of Blockchain as it provides transparency in transactions. The usage of CNN and the Xgboost regressor model showed maximum performance. The implementations of data cleaning, data preprocessing and model implementation to give precise, sound, and consistent data were executed successfully. The results were noted and the corresponding graph is plotted.

In this paper, we have concentrated only on the scrapping of cars. In the future, it can be diversified to scrapping of all vehicles. Moreover, to bring out a meticulous classification model, damage detection can be applied to a larger dataset. The system can also be diversified in the future to includenew modules such as recycling and selling of scrapped parts.

VII. References

[1] https://pib.gov.in/PressReleaseIframePage.aspx?PRID=1705811
[2] Toqeer Ali Syed, Muhammad Shoaib Siddique, (Member, IEEE), Adnan Nadeem, (Member, IEEE), Ali Alzahrani, Salman Jan and Muazzam A. Khan Khattak, A Novel Blockchain-Based Framework for Vehicle Life Cycle Tracking: An End-to-End Solution, IEEE Access (Volume: 8)
[3] Mehdi Benchoufi , Raphael Porcher , Philippe Ravaud, Blockchain protocols in clinical trials: Transparency and traceability of consent, https://www.researchgate.net/publication/318200735
[4] Lin William Cong. Zhiguo He, Blockchain Disruption and Smart Con-tracts, https://www.researchgate.net/publication/333388218 – smart con-tract
[5] [1] Tongzhu ZHANG, Xueping WANG, Jiangwei CHUI, Xianghai LIU, Pengfei CUll, Automotive Recycling Information management based on IoT and RFID
[6] Jie Li1, Michael Barwood1 and Shahin Rahimifard1, An Automated Approach for Disassembly and Recycling of Electric Vehicle Components
[7] [3] Mohammad Z.Masoud, Yousef Jaradat, Ismael Jannoud and Dema Zaidan, CarChain: A Novel Public Blockchain-based Used Motor Vehicle History Reporting System
[8] Samiksha Marne, Shweta Churi, Delisa Correia, Joanne Gomes, Predict- ing price of cryptocurrency – A deep learning approach, NTASU - 2020 (VOLUME 09 - ISSUE 03)
[9] Rishabh Ranka, Niranjan Sharma, Naman Talati, Nitika Rai, An Efficient System for Implementation of Goods and Service Tax in India using Blockchain, NTASU - 2020 (volume 09 - issue 03)

[10] Shriram. S, Dr. K. Anuradha, Dr. K. P. Uma, Future Stock Price Prediction using Recurrent Neural Network, LSTM and Machine Learning, ICRADL - 2021 (Volume 09 - issue 05)
[11] Devdoot Maji , Ravi Singh Lamkoti , Hitesh Shetty , Bharati Gondhalekar, Certificate Verification using Blockchain and Generation of Transcript, volume 10, issue 03 (March 2021)
[12] Jianxin Wu, Introduction to Convolutional Neural Networks, https://cs.nju.edu.cn/wujx/paper/CNN.pdf
[13] Tianqi Chen,Carlos Guestrin, XGBoost: A Scalable Tree Boosting System, https://arxiv.org/pdf/1603.02754.pdf