Department of Electrical Engineering

FINAL YEAR PROJECT REPORT

BENGEGU4-INFE-2019/20-YNS-04

Compare the performance of SVM and a deep learning model for sequence classification

Student Name: Wong Ka Ho
Student ID: 54816829
Supervisor: Dr SUN, Yanni
Assessor: Dr CHAN, Rosa H M

Bachelor of Engineering (Honours) in

Information Engineering

Student Final Year Project Declaration

I have read the student handbook and I understand the meaning of academic dishonesty, in particular plagiarism and collusion. I declare that the work submitted for the final year project does not involve academic dishonesty. I give permission for my final year project work to be electronically scanned and if found to involve academic dishonesty, I am aware of the consequences as stated in the Student Handbook.

Project title : Compare the performance of SVM and a deep learning model for sequence classification

Student Name : Wong ka ho

Student ID: 54816829

Signature Issac

Date : 15/11/2021

Abstract

In recent years, DNA sequence classification is a crucial challenge, because viruses and bacteria seriously damage daily life, public health, and the economy. Classifying the harmful virus before the happening global pandemic disease is a good method to prevent the next serious impact caused by viruses like COVID-19. Meanwhile, Several machine learning techniques, which include deep learning and supporting vector machine(SVM), have succeeded completed the task. In this research, we complete the performance of deep learning and supporting vector machine by different criteria. The experimental result shows that the performance of deep learning in sequence classification is better than supporting vector machines.

Acknowledgement

I want to thank my project supervisor Dr. Yanni Sun and accessor Dr. Rosa Ma. They assign me a very interesting research topic. I can explore the machine learning area. Having a chance to study the knowledge and have hands-on experience in machine learning modeling.

In addition, I also thank you my student mentor for this project Xubo. He gives lots of useful advice in my experiment. He teaches me many useful skills in searching machine learning resources. He also guides me on how to construct two machine learning models and how to do the data preprocessing patiently. He clarifies my mistake and fault with the machine learning model and gives me some useful suggestions to improve. Many thanks for the kindly suggestion of Xubo.

I also appreciate the contribution of technical staff work with electrical engineering high-performance computer center (eehpc). They provide a very stable high-performance computer to students. Let us can connect the high-performance computer conveniently. We can train and test our machine learning model on the computer with less time-consuming.

Worth mentioning, I am very grateful to join EE department. I learn a lot of knowledge of electrical engineering from here. All the EE department staff is very helpful and kind. They provide much help to me when I feel difficulty in studying EE courses. Especially Ms. Stephanie Wan, She gave me much encouragement when I feel depressed. Many thanks to all the helpful EE staff who give me continuous help in my student life.

Table of contact

List of figures

List of table

Chapter 1

1. Introduction

1.1 background

There are 1.6 million viruses in the world, as well as viruses such as COVID-19, which jumped out of mammals and humans. The author has studied in detail the impact of COVID-19 in many areas. Because of the impact of globalization and the strong mobilization of the global people, new viruses are easily emerged and spread. For example, COVID-19 spread fast and become a global pandemic. Early identification of viruses, fungi, and bacteria can help prevent outbreaks and speed vaccine design. Thus, the classification of DNA sequences acts as an important role in computational biology.

Machine learning is a powerful technique for analyzing large-scale data, which can be learned spontaneously to acquire knowledge. It can be widely used for analyzing the DNA sequence. Moreover, lots of the research result has proven the result of machine learning in analyzing the DNA sequence.

1.2 Motivation

The application of AI has slowly penetrated our daily lives, doing more and more things for us. The development of Artificial intelligence become mature. Many different branches and the technical principles are also diverse. The hottest deep learning & SVM are introduced here.

Different from traditional software, the main idea of deep learning is to simulate the brain of humans. Traditional software is a tool to help people finish tasks more efficiently. It can simplify the solution to dealing with the task. In deep learning, it is trying to build artificial intelligence. The deep learning model can make a decision like a human after training with a certain dataset. The machine can predict the best result by considering the rules set by the user. Thus, it can apply the knowledge in the real world to solve real problems. The deep learning model does not like the traditional software to provide exact solutions to different problems. Deep learning is closer to human learning, the model needs to study the more different cases and catch the common feature and then make a conclusion. Therefore, it is more abstract and harder to understand.

Support vector machine (SVM) which is a binary classification model. The basic model of the support vector machine is a linear classifier. Its main target is to define the maximum interval in the character area. Different from other machine learning algorithms, support vector machine also includes kernel technique. It can make a change of support vector machine to be a non-linear classifier. The main strategy of the support vector machine is maximizing the interval of data. It can solve the convex linear problem with formalization. It

is also minimizing the regularized loss in the problem. Therefore, a support vector machine can have a good performance in solving the convex linear problem.

1.3 Objective

In this project, We will construct the deep learning model and support vector machine model to abstract the high-level character from minimum preprocessed data. In this research, we will classify DNA sequences using convolutional neural networks and supporting vector machines while treating these sequences as text data. I transform the DNA sequence from textual to number by ordinal encoding. The arrangement of nucleotides determines the character of DNA. Thus, the basic position of each nucleotide will not be changed. The transformed data will be an input layer of the model. The deep learning model and support vector machine model will be trained with the training dataset and test with the testing dataset. We will evaluate our models and compare two model performances and achieve significant improvements on all these data sets.

Chapter 2

2. Literature review

2.1 Sequence classification

Sequence refers to a contiguous memory space that can store multiple values. These values are arranged in a certain order, and they can be accessed through the number (called index) where each value is located. In order to understand the sequence more vividly, it can be regarded as a hotel, then each room in the store is like a memory space for storing data in the sequence, and the unique room number of each room is equivalent to the index value. In other words, we can find each room (memory space) in this hotel (sequence) through the room number (index).[1]In Python, sequence types include strings, lists, tuples, sets, and dictionaries. These sequences support the following general operations, but what is more special is that sets and dictionaries do not support indexing, slicing, addition, and multiplication operations. Sequence Classification order is responsible for estimating categories according to surveillance programs. In many applications, such as healthcare surveillance or intrusion detection, early identification is essential for timely intervention. In this job, we learn a sequence classification that helps you change from a learning path to an early stage. The most advanced sequence classification works in neural networks, especially for LSTMs, we classify finite-state automata forms and study them through discrete optimization. Our automata-based classification provides definitions, definitions, and definitions. They have a strong practical capacity for reasoning and the improvement of the human cycle. Experiments on target recognition and behavioral classification data establish

that the automata-based classification we studied is more robust and comparable to the LSTM [2]-based classification

2.2 SVM

Support vector machine(SVM) is a kind of machine learning algorithm. It mainly applies in the situation that we need to classify them belong to different classes.[1] Then we need to find a linear classifier that can separate the data into two sets. The strategy of the linear classifier is by considering the feature of the linear combination to predict the result, but it can not determinate by the feature of the non-linear calculation result.

For example, they are a set of data with two different features (fig 2.2.a). We need to use a line to classify them into two classes(Fig2.2.b). Similarly, if we can separate the data with the surface in three-dimensions space.

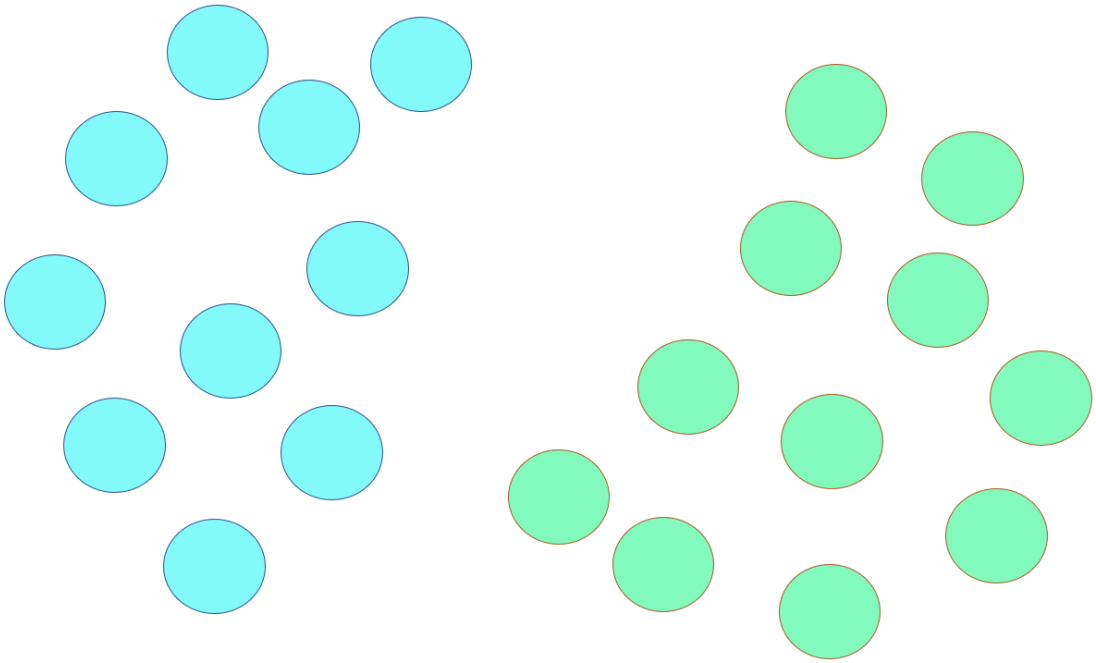

Fig 2.2.a Dataset of SVM

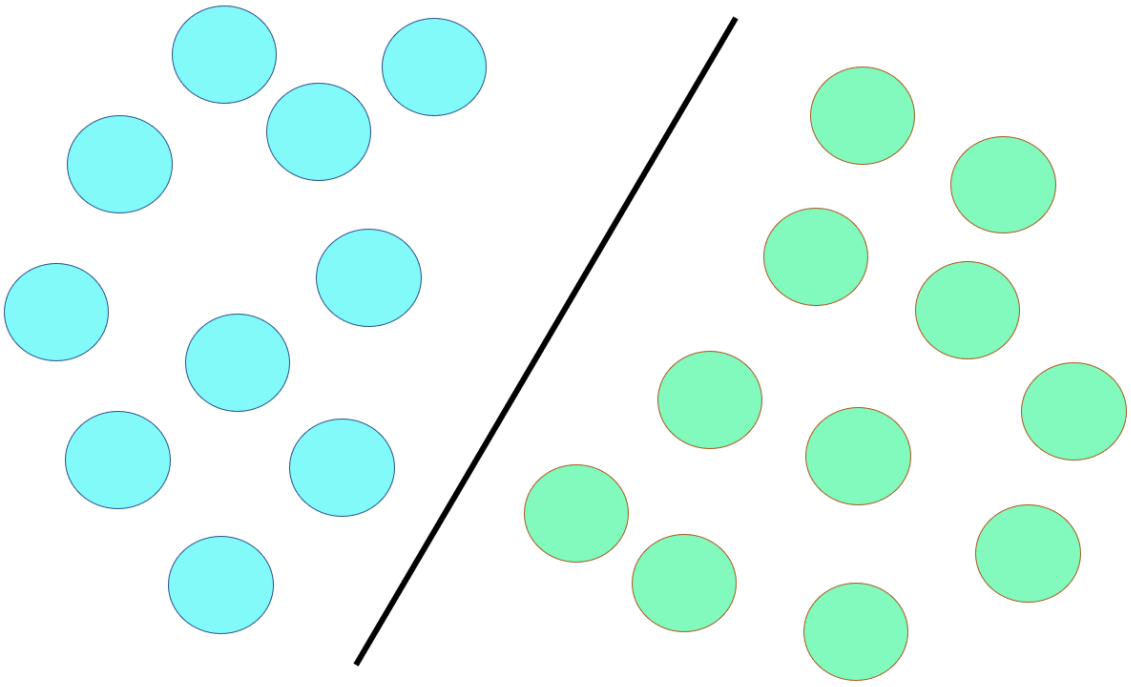

Fig 2.2.b Classifies two datasets

Support vector machine is a good algorithm to find the line to classify data into two datasets. The target of the support vector machine is to find the best line to classify the data into two classes. The working process of the support vector machine is that the support vector machine will set a line (separation line) to divide the dataset into two classes.[2] Then, the Support vector machine will create two parallel lines of the separation line. The interval between the separation is called margin(fig 2.2.c). The strategy of the support vector machine is to find the best line by maximizing the margin interval.

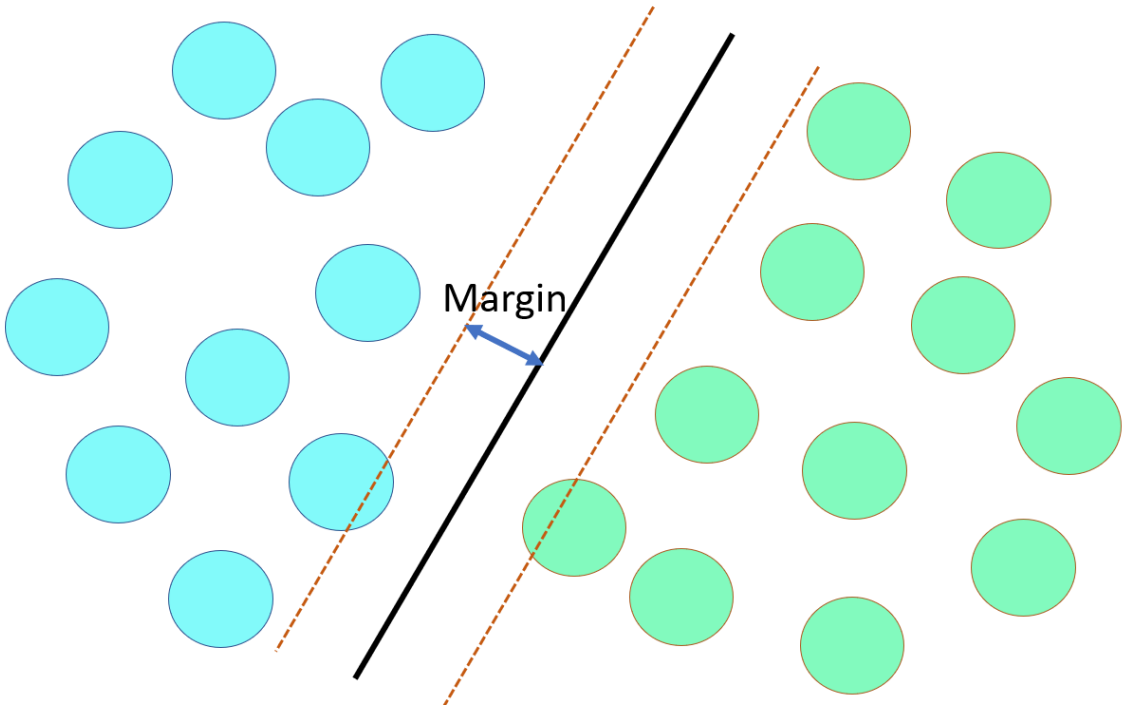

Fig 2.2.c Margin of SVM

2.2.1 The algorithm of SVM

To maximize the margin, we need to make a few assumptions. Let the separation line be x*y = 0, Thus, the upper area will be x*y>0 and the lower area will be x*y <0.(Fig 2.2.1.d)

Similarly, the line in the upper area is x*y = -k and the line in the lower area will be x*y = k.

Therefore, the area upper than x*y = -k will be x*y < -k, and the area lower than x*y = k will be x*y>k(Fig 2.2.1.e). Remind that the area between the dotted line will not have any data.[3]

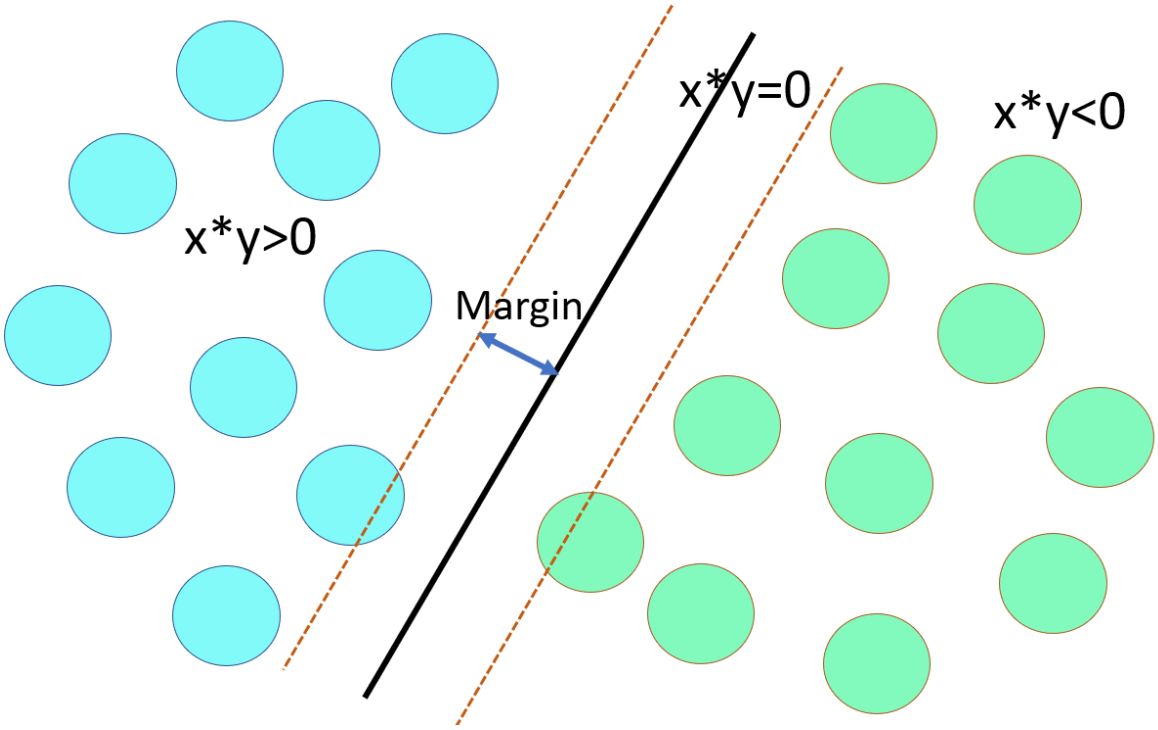

Fig 2.2.1.d Area divide by separation line

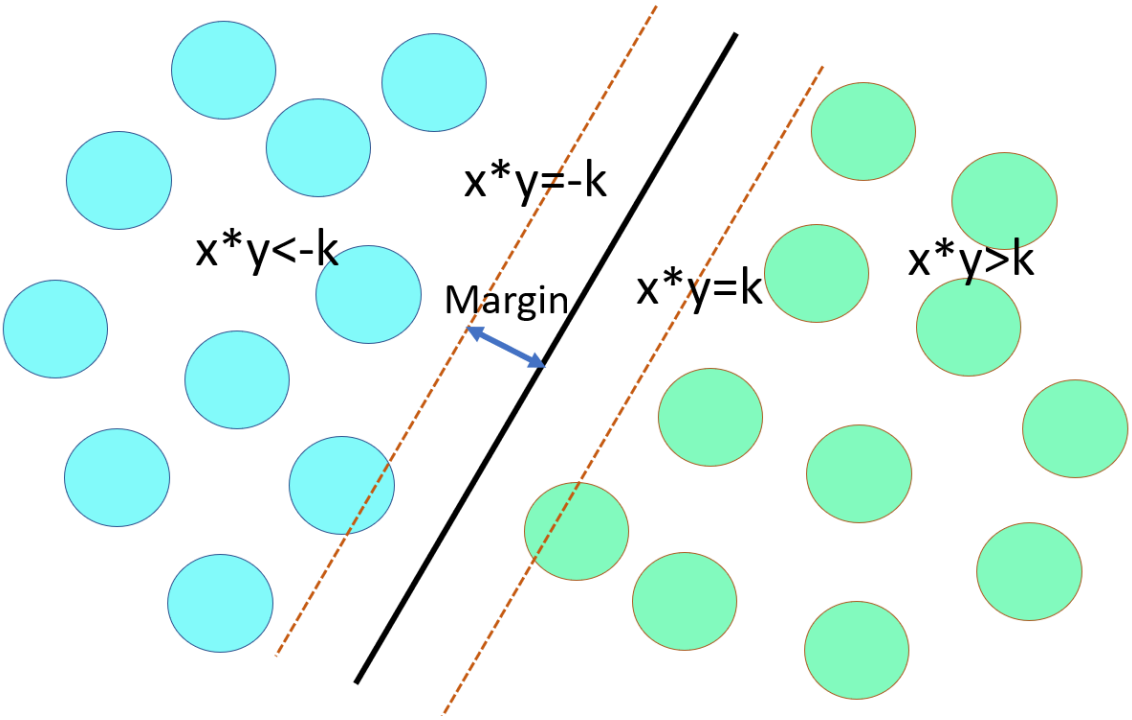

Fig 2.2.1.e Area separated by dotted line

In this situation, we will use a support vector to calculate the margin and then maximize it. First, we need to create two new vectors X1 and X2. Then we need to find a projection vector W by project X1-X2(Fig 2.2.1.f). Finally, we can maximize the margin by considering the equation Y*(W*X).[4]

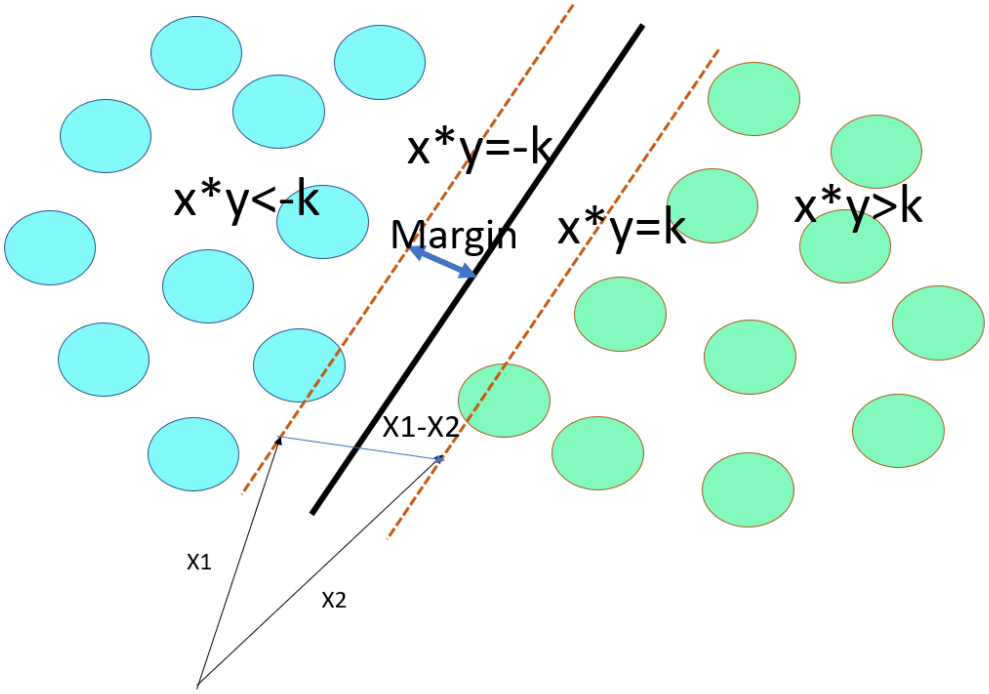

Fig 2.2.1.f Projection vector

2.3 deep learning

Deep learning is a type of machine learning. The main idea of the deep learning algorithm is to simulate the human brain to achieve human thinking in the machine. Before we understand the deep learning algorithm, we need to learn some basic knowledge of the human brain.

The human brain is constructed by a large number of neurons. For finishing any task, all the neurons will cooperate to run and transfer information. Form human catch the information received in outside, handle with brain and send a corresponding action. The information is passed through many neurons within a few seconds. All the neurons are training continuously. Difference neurons will execute their own job when they received the information. Some of them are for transferring, some of them are analyzing, some of them are

for storing. [5] Thus, the performance of the human brain depends on the knowledge of the brain.

By considering the structure of the human brain, the target of deep learning is to create a much different calculation layer to simulate a large number of neurons and train with big data to achieve human thinking in the computer. Although the result of deep learning cannot be found 100% accuracy, the functional relationship with the output can be as close to the actual relationship as possible. [6] The deep learning model will analyze the data in every designed layer by considering the feature of the dataset. After a certain training time of the model, the data can be classified or predicted by the model easier. [7] In each layer will have a core mission and have certain parameters. However, the parameter should be controlled in a certain range. Otherwise, it will make the model overfitting. The result of overfitting will lead the model to only have good performance in the training data.[8]

2.3.1 Convolutional Neural Network Model

Convolutional Neural Network(CNN) is one of the most respected algorithms of deep learning. In our research, we will choose convoluted neural networks to represent deep learning to perform sequence classification.

The convolutional neural network is inspired by the structure of the visual system. The first convolutional neural network computing model was proposed in Fukushima's neurocognitive machine.[9]

The convolution neural network is mainly divided into two parts: feature learning(Fig 2.3.1.a) and classification(Fig 2.3.1.b).

The feature learning part includes the convolution layer, pooling layer.

The classification part includes a flatten and a fully connected layer.

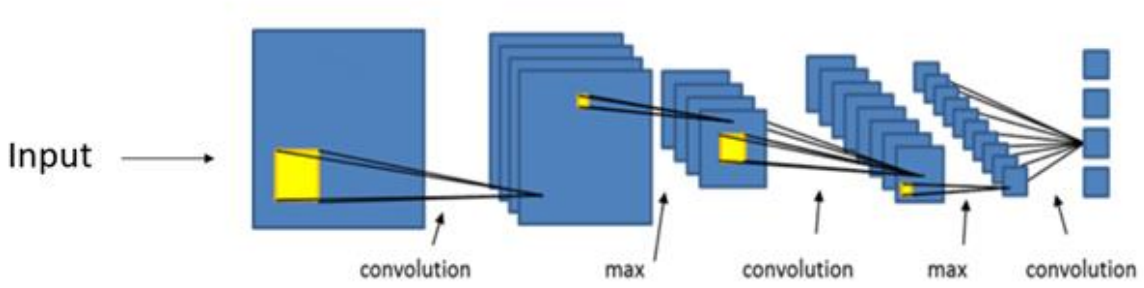

Fig 2.3.1.a Feature learning of CNN

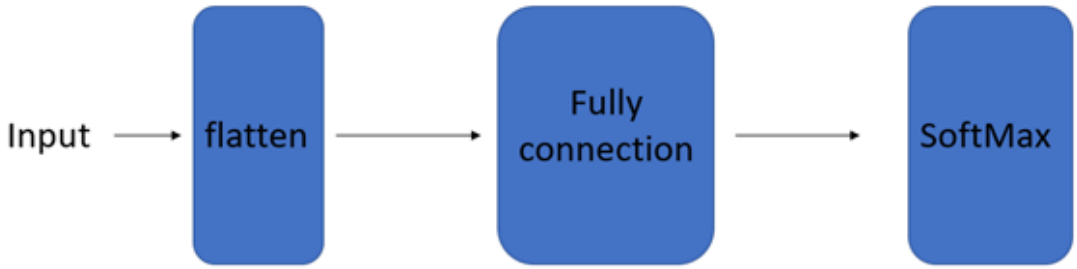

Fig 2.3.1.b Classification of CNN

convolution layer

The principle of convolution layer is fetching the input layer data feature through rolling the convolution kernel. Output the feature map by padding and stride controlling of the convolution kernel.[10]

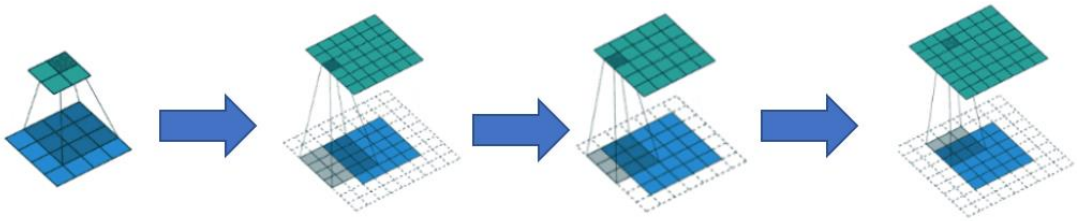

Fig 2.3.1.c Example of convolution layer

Pooling layer

The main advantage of the pooling layer is that decreasing the complexity while increasing the efficiency of calculation through compressing the input layer data. Different from the convolution layer, the pooling layer does not have parameters. It only calculates the maximum and the average of a vector.[11]

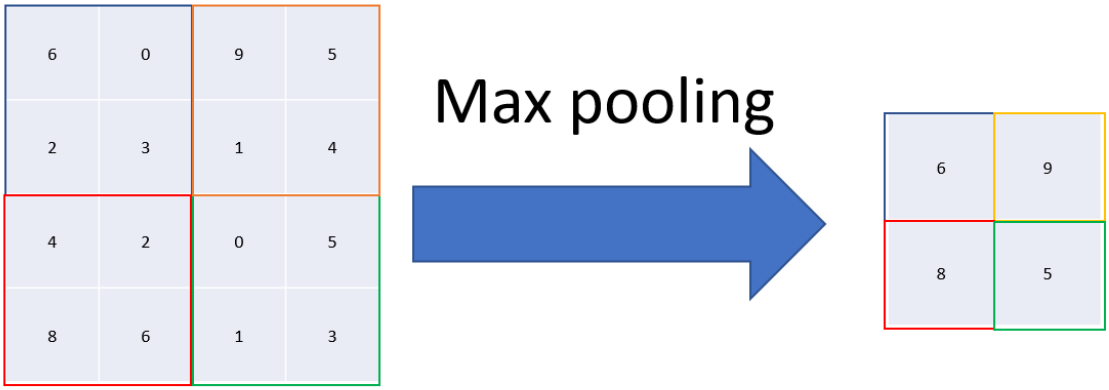

Fig 2.3.1.d Example of pooling layer

Flatten

Flatten is mainly acting as a transition layer between the convolution layer and fully connected layer. Flatten will compress the n-dimensional matrix to a one-dimensional matrix and output it.[12]

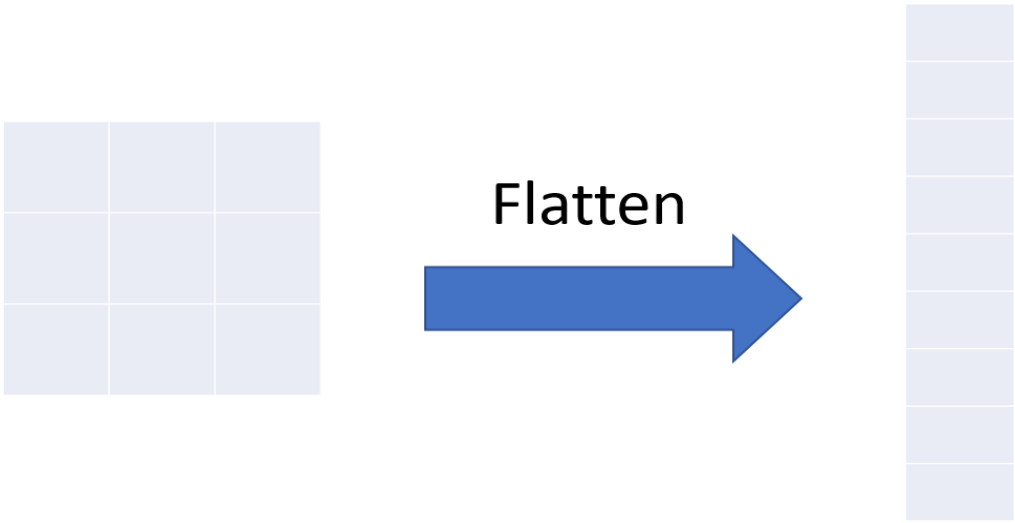

Fig 2.3.1.e Example of flatten

Fully connected layer

The most important usage of the fully connected layer is classification. In this layer, the feature of the data will be fetched. By considering the fetched feature and calculating their weighting, the data will be classified. [13]

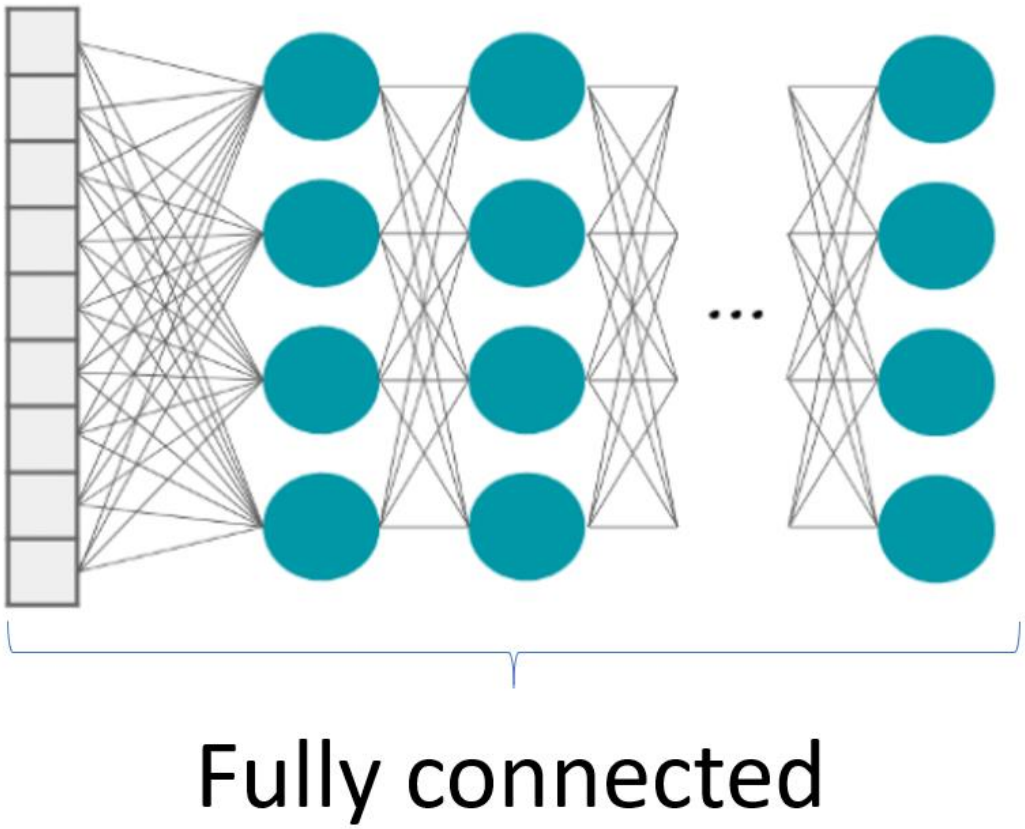

Fig 2.3.1.f Example of fully connected layer

2.4 different between SVM and deep learning

Generally, support vector machine can finish some tasks but deep learning cannot be. Meanwhile, deep learning can perform well in some tasks but support vector machines cannot. Many deep learning algorithms can be regarded as special cases. For example, a shallow neural network is a special case of a deep learning algorithm. Thus, we choose a convolution neural network to represent deep learning. The biggest difference is nothing more than that the model of deep learning is more complex, which derives a larger model and algorithm system, thereby achieving greater progress in functionality.[14]

The algorithm structure of deep learning can be more flexible in model design than support vector machines. All differentiable arithmetic processes can be added to the network structure of deep learning. For example, the soft thresholding commonly used in noise reduction algorithms is differentiable.[15] Embedding the soft threshold into the classic residual network, a deep residual shrinkage network suitable for strong noisy data is obtained. Moreover, this structure allows end-to-end training. This is a function that traditional machine learning algorithms represented by SVM do not have.[16]

In the perspective of traditional statistics, the theory of convolution neurall network is based on the sample data is infinitely large. That means the convolution neural network model is the statical properties when the sample data tends to infinity. However, we are difficult to get an infinitely large data sample. We only can train the model with a limited dataset. Thus, it is difficult to achieve the desired derivation result. In contrast, support vector machines can overcome the inevitable problems of neural networks. Due to the fact that support vector machines have stronger approximation power and generalization ability.[17]

Chapter 3

3. Detailed Methodology and Implementation

In this chapter, we will provide detailed information about our model in the test analysis data. In addition, we demonstrated how to implement and compare these models or algorithms in our experiments. Overall, it outlines how we conduct the test.

We adopted deep learning and SVM as the basic model, whereas these two-machine learning algorithm and structure are proposed as comparative models in our experimental analysis.

3.1 sequence representations

Each of us shares our air, food, water and shelter with tiny microflora including viruses, bacteria and fungi. Most of these tiny microorganisms are harmless, but some are pathogens—the kind that can make you sick, such as the new coronavirus that causes COVID-19.

Deoxyribonucleic acid(DNA) is an essential biological macromolecule for the normal functioning and development of organisms. Every organism must contain DNA. DNA is mainly combined with four nucleotides which are Adenine (A), cytosine (C), guanine (G), and thymine (T).(Fig 3.1.a) The properties and the categories of the DNA are determined by the arrangement of the four main nucleotides. Each of the fungi, bacteria, and virus have a

unique nucleotides arrangement. Thus, the arrangement of DNA is like an identity card of fungi, bacteria, and virus.[19]

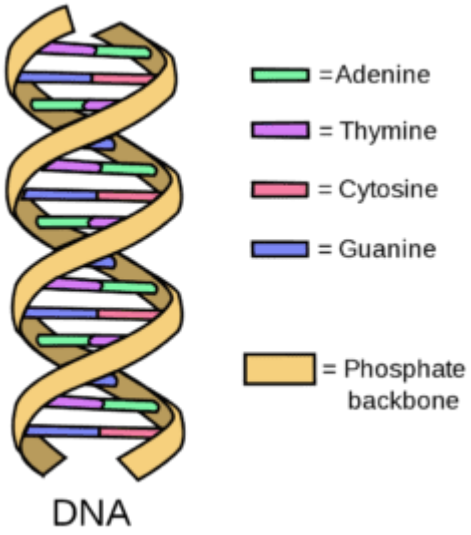

Fig 3.1.a Structure of DNA

3.2 Data Preprocessing

Due to the fact that the data is incomplete in the real world and the machine does not know the human language. The machine only knows “machine language”. Thus, we need to do data preprocessing to convert textual data to numerical data. The goal of data preprocessing is to translate the data that the kernel can understand through the programming method. [20]In data processing and feature engineering, we often encounter type data. Converting numerical value and constructing the data as an input layer is very essential, but the model often defaults to continuous numerical processing. It will lead to the bad performance of the model and increase the difficulties in the model designation. Meanwhile, Encoding is a good method to solve this problem. The encoding method is to use different numbers of registers to encode

different states. Each state has its independent number in every moment within the training. Every number will use only one time. The advantages of ordinal encoding are the ability to handle non-continuous numerical features and the features are also expanded to a certain extent.[21] In the following, we will illustrate the steps of data processing that we can perform on DNA sequence text data.

In order to build a deep learning model, I created a dictionary of 4 DNAs. The dictionary is further used for integer encoding in increasing order. And it is a kind of ordinal encoding method. The reason why use ordinal encoding but not one-hot encoding is ordinal encoding has better performance than one-hot encoding. Using ordinal encoding can reduce the input dimensionality and running time of the input layer in CNN but the fetching rate of the CNN model is dependent on the layer structure.[22]

original data	ordinal encoding
bad	1
good	2
excellent	3

Fig 3.2.a Example of ordinal encoding

3.3 SVM implementation

3.3.1 flow of SVM

After the data preprocessing, we will split the data into two set: training data and testing data. we can use the support vector machine (SVM) as the main algorithm. The support vector machine model will be trained by a training dataset. In addition, we will use the support vector machine classifier(SVC) to act as the model classifier to classify the test data and generate the result. Those program functions are exported from sklearn and Keras.

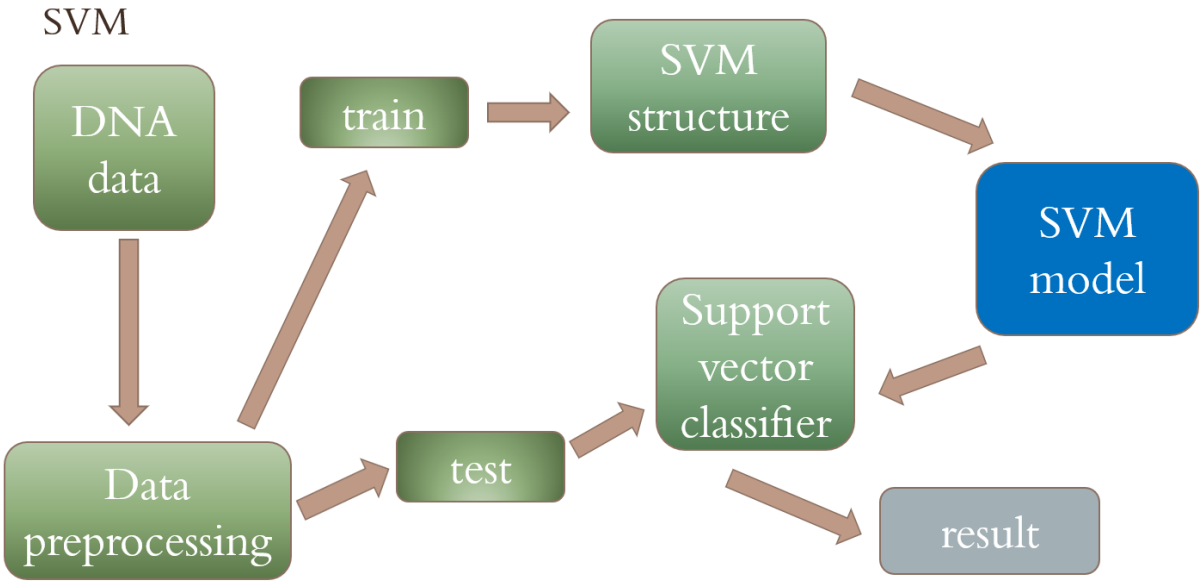

Fig 3.3.1.a Flowchart of constructing SVM model

3.3.2 Support vector machine classifier

Support vector machine classifier (SVC) aims to predict a continuous output value. For example, suppose you are trying to predict the revenue of a certain brand based on many input parameters. The regression model is actually a function that can output potentially any

income figures based on certain inputs. It can even output income figures that have never appeared anywhere in the training set.[23](Fig 3.3.2.a)

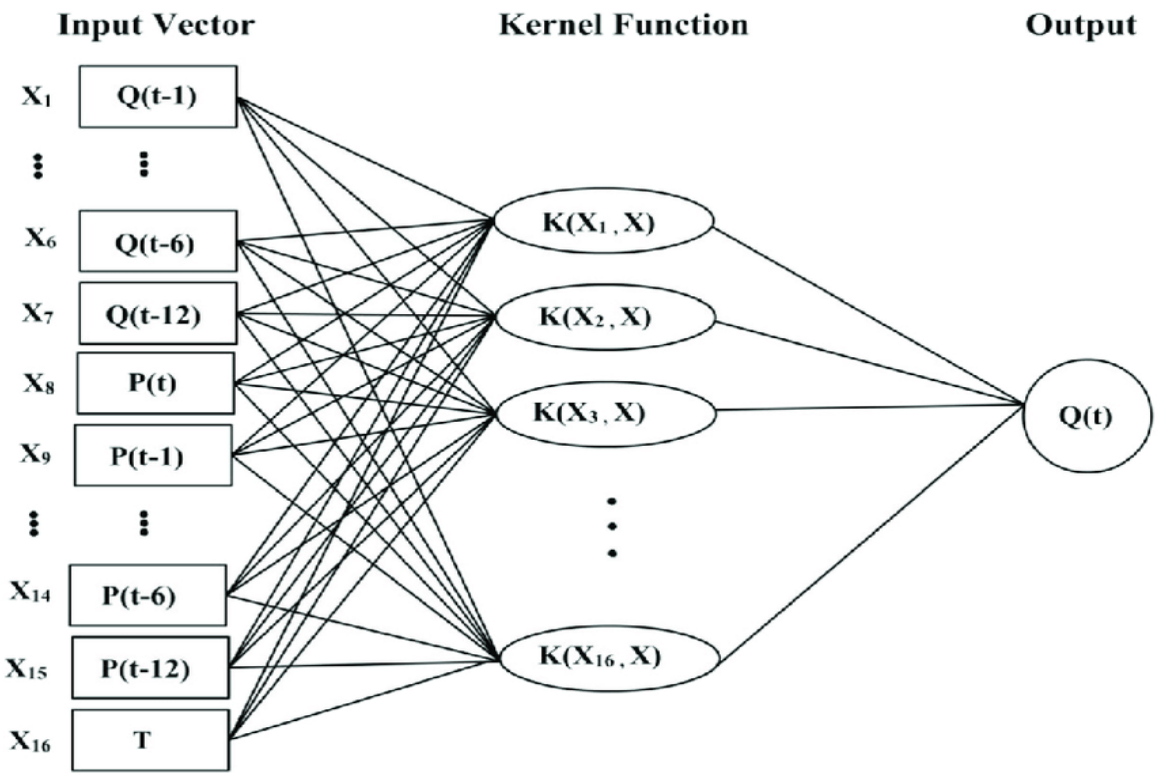

Fig 3.3.2.a The structure of SVC

3.3.2 Aims of SVM model

Classification aims to predict which class (discrete integer or classification label) the input corresponds to.[24] For example, suppose we have divided sales into low sales and high sales, and you are trying to build a model (binary/two-level classification) that can predict low or high sales. The input may even be the same as before, but the output will be different. In the case of classification, your model will output "low" or "high", and theoretically, only one of these two responses will be generated for each input.

3.4 deep learning model implementation

Deep learning is a very large category. For the following model training, we will adopt the convolution natural network(CNN) mainly to analyze the performance. Under convolution natural network, ResNet is used for build the CNN architecture.

3.4.1 Deep residual network(ResNet)

Deep residual network has very strong expression power in classification and it can reduce the workload in fetching features. The advantage is of a deep residual network is that it can control the network degradation problem.

The network degradation is a challenge in convolution neural networks. The performance of CNN will be better even tends to be perfect but will be worse speedy suddently with increasing the depth of CNN. The details of the testing error and training error is shown in the below figure.[25]

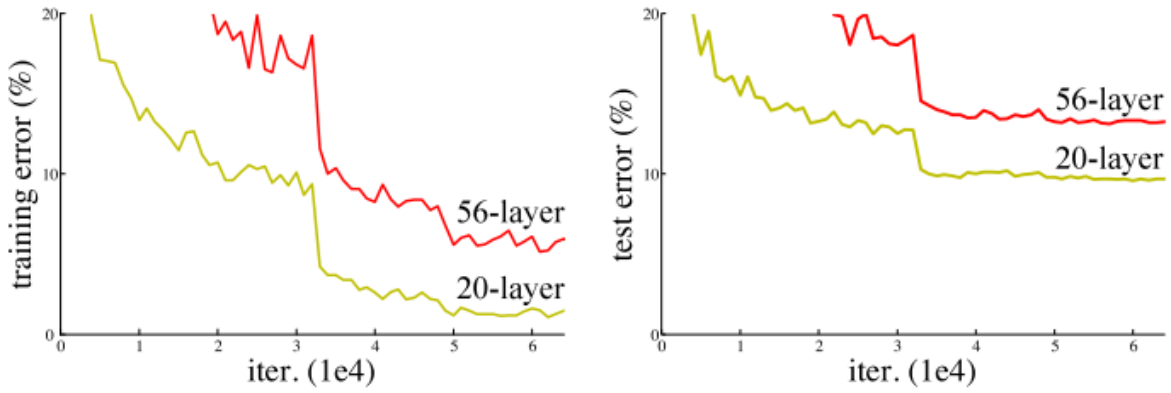

Fig 3.4.1.a Figure of network degradation

To solve this problem, ResNet creates a residual block to achieve jump layer connection by assuming the dimension of input and output are equal. The main structure of a residual block activates the summation of input unit and output unit.[26]The simple structure of a residual block is shown below.(Fig 3.4.1.b)

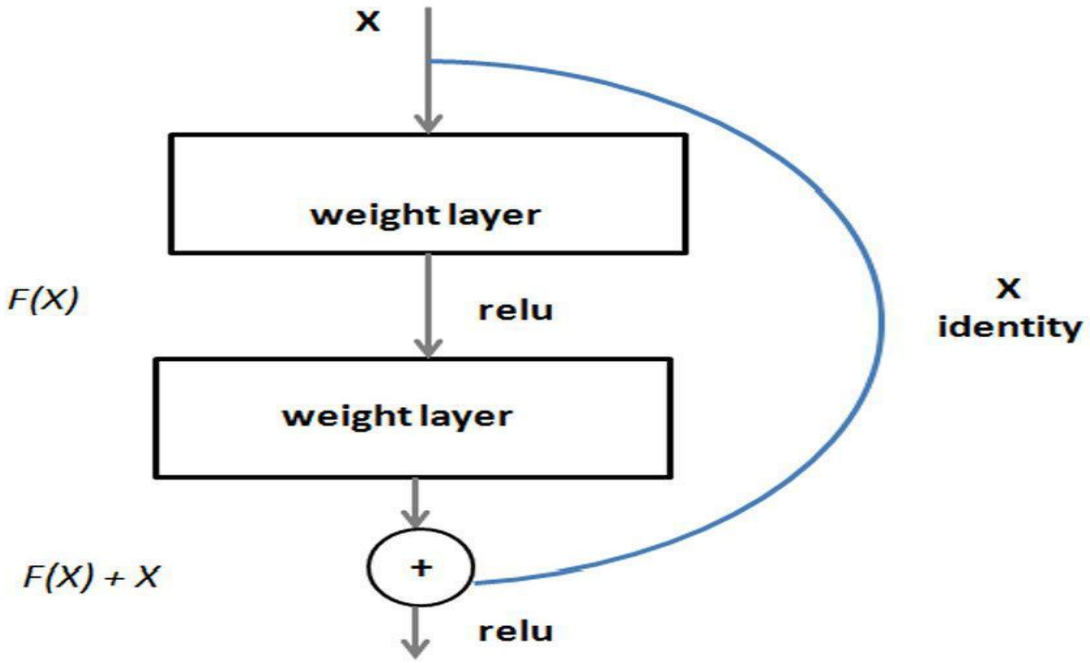

Fig 3.4.1.b Simple structure of residual block

3.4.2 Deep residual network Architecture

The DNA sequence is converted into a hot code with a shape of (batch_size, 100, 5) as the input of the model. The initial convolution layer will be one-dimensional convolution layer with zero padding. In convolution layer, the feature of the data will be fetched by the filter. The convolution of the feature will be send to next layer. Then, the residual block will

capture the output from input layer and process the data inside the residual block. The data will output to pooling layer after two residual blocks. The pooling layer can decrease the calculation workload of the model by calculating the weighting of the parameter and feature. Thus, the pooling layer also can decrease the usage of memory. Then, the data will be flatten. The data will convert to one dimension and past to dense layer. The dense layer also know as fully connected layer. It can re-fitting the feature to reduce the loss of the model result.(Fig 3.4.2.a)

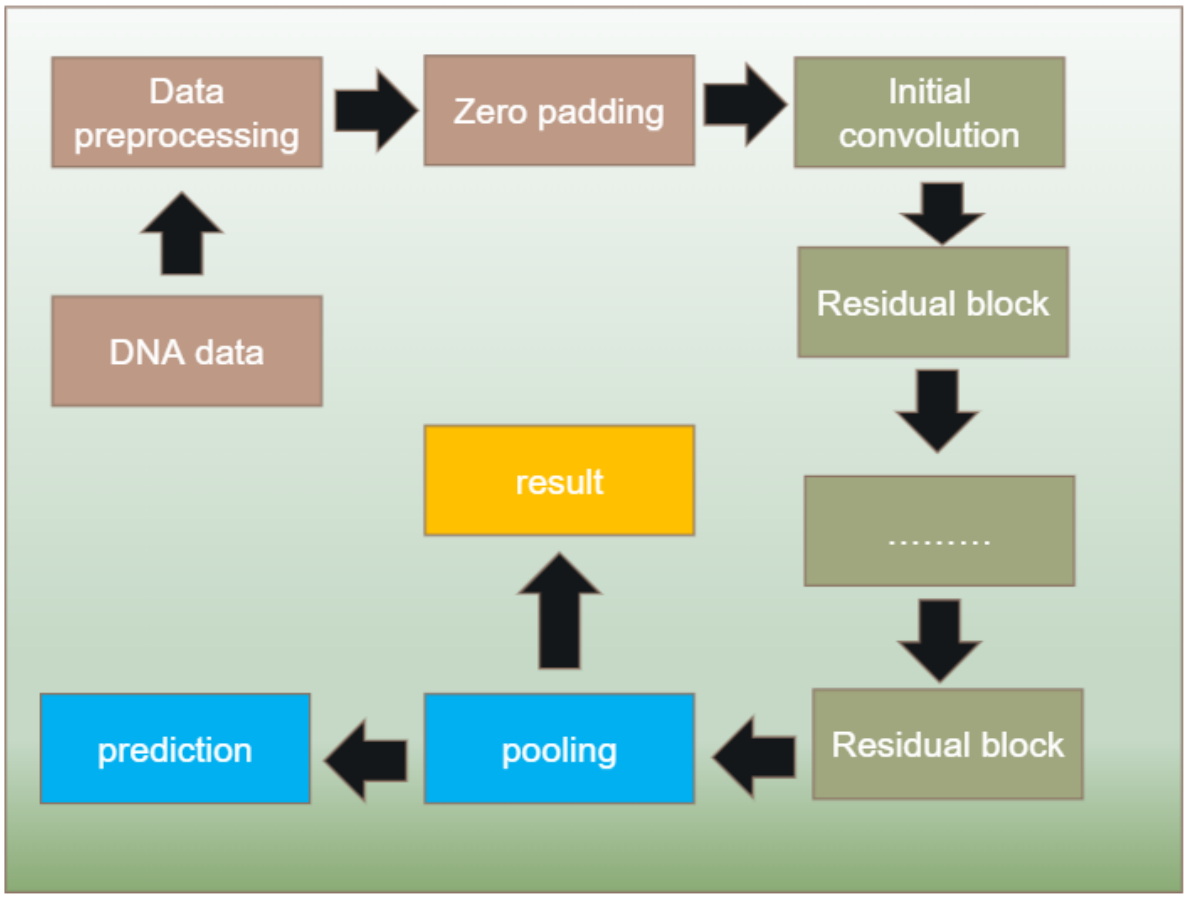

Fig 3.4.2.a Structure of CNN model

The detail structure(Fig 3.4.2.b) of residual block is that the data will pass through Batchnormalization layer then Relu layer and convolution layer. The cycle will perform twice. In the first convolution layer, the kernel size will be 1x1.The kernel size in the second convolution layer will be 3x3.

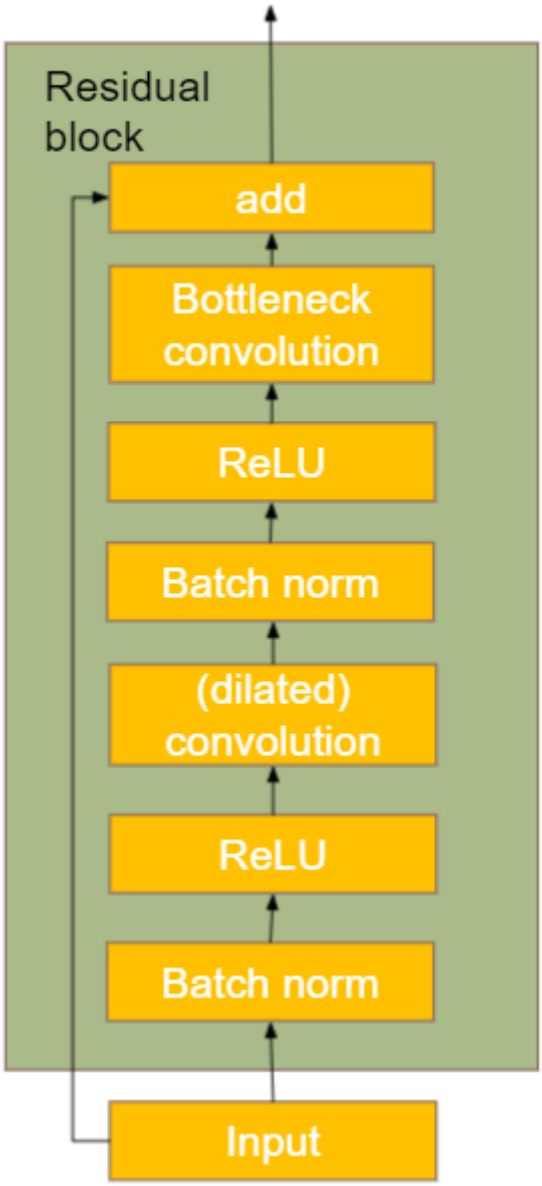

Fig 3.4.2.b Structure of residual block

3.5 Methodology and steps to run a test

As the aims of the experiment is compare the performance of SVM and a deep learning model for sequence classification, there are three main topic: sequence classification, SVM, deep learning. For building a good experimental environment and running the above-mentioned approach, We generalize a few steps as follows.

i) Prepare data set: search and download the data about sequence classification. We prepared a set of DNA data that mixed bacteria, fungi, and viruses.

ii) Prepare computer: Prepare a more powerful computer to perform machine learning.

iii) Load the data: upload the data to the prepared computer and load the DNA dataset.

iv) Text preprocessing: converting these text data into digital forms that can be processed by the machine for building the model.

v) Create two models: construct the algorithm and structure of deep-learning and SVM model respectively

vi) Define comparing platform: compare deep-learning and SVM model with accuracy, recall, precision, and F1 scores.

vii) Train the model: fit the data into SVM and deep-learning model and train two models.

viii) Save and load models: Save the model weight during and after training. The model can be restored from where it stopped which is aims for avoiding long training.

ix) Test the model: testing two models with test data. Evaluate and predict the sequence classification performance of two models.

x) Analysis of the result: Recording the result of each model in a text file. Tabulate the result data with excel for a better analyzing environment.

Chapter 4

4. Experimental Study

In this chapter, we will illustrate the experiment environment and experiment setup process. After getting the knowledge of the two machine learning algorithms and by considering the step mentioned before, we will try to build two models and compare their performance.

4.1 Experimental Setup

4.1.1 Hardware

Due to Machine learning is a very consuming computing resource. GPU is the heart of machine learning. A good GPU can increase the speed of the training process significantly. Thus, I borrow a high-performance computer in the electrical department. We will use the computing node in the high-performance computer to implement two models. The detailed information of the computing node is shown below(Table 4.1.1.a)

Computing node	Information
	2 X AMD EPYC 7522 64-core processors (128 cores/node)
	512 GB Memory (Useable memory space ~ 448GB)
	CentOS 8.1
	400GB Local storage at /local

Table 4.1.1.a Detail information of computing node

4.1.2 Software

Within the construct two machine learning models, I adopt some popular open-source tools to help me program two machine learning models.

1. In the operating system, I use Linux. The machine learning development environment relies on the GPU, and the Linux configuration speed of GPU is faster than Windows and macOS Linux is a command-based operating system. It can perform high stability with low-cost hardware.

2. In a compiler, I use Jupyter notebook. It is an open-source web application. It contain a large amount of useful library. It provides an environment that user can write a brief introduction to the program in a very clear format . In view of these advantages, it can help me to build a machine learning conveniently, such as data preprocessing, statistical modeling, training and testing machine learning models, etc.

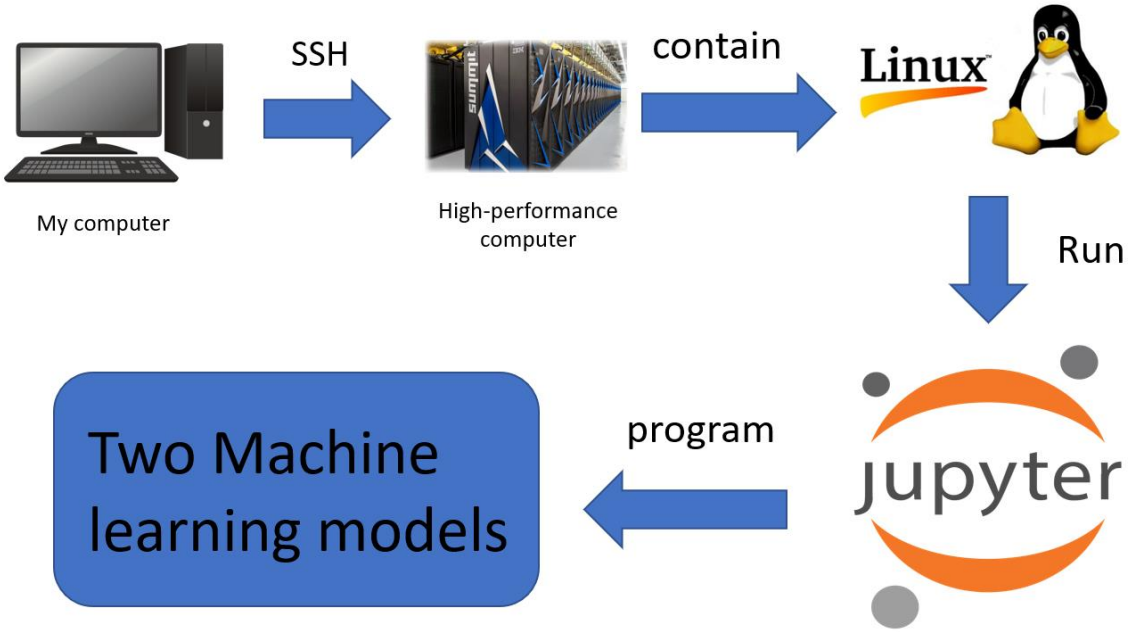

Fig 4.1.2.a Overview of programming

4.2 predictions of the experiment

Before I start the experiment, I do some guess and prediction in the performance of support vector machine and convolution neural network in sequence classification.

1. Convolution neural network is an algorithm that simulate the human brain to perform sequence classification. The structure of human brain is very complicate compare with normal CPU. Hence, I guess the training time of Convolution neural network will be longer than support vector machine.

2. In calculation or big data management, the performance of computer is better than human brain. As mentioned before the design idea of convolution neural network is simulate human brain. Hence, I guess the accuracy of support vector machine will be better the convolution neural network.

3. The programming structure of convolution neural network is more complicate than support vector machine. Convolution neural network need to build up many layers and more data preprocessing. Hence, I guess the model performance of convolution neural network will be better than support vector machine.

4.3 Datasets and experimental result

In this section, I am going to present the details of the experimental dataset and result. The dataset is a DNA sequence mixed with fungi, virus and bacteria. Each DNA sequence contain unique arrangement. Every DNA sequence are labeled with 0,1 or 2. 0 means virus, 1 means fungus, 2 means bacteria. All the data is store at csv format.

In convolution neural network model, The data set is split to 3 part: validation data, testing data and training data. The size of each data is shown below.

data	size
train	104857
val	170666
test	41210

Table 4.3.a CNN data size

In support vector machine model, The dataset is split into two part testing data and training data. The size of each data is shown below.

data	size
train	104857
test	10485

Table 4.3.b SVM data size

After train and test the model by prepared dataset, the result has generated in below. In the result I will mainly use 6 standard to analyze it.

SVM	
Standard	scores
Accuracy	0.754760675
Precision	0.834042549
Recall	0.502816878
F1 score	0.501067857
Time	Near 24 hours

Table 4.3.c Result of SVM

CNN	
standard	scores
accuracy	0.867328653
precision	0.753968479
recall	0.662816878
F1 score	0.681067857
Time	Near 3hour

Table 4.3.d Result of CNN

Chapter 5

4. Discussion

In this chapter, I will analyze the result by form 6 standard and compare the performance of convolution neural network and supporting vector machine in sequence classification.

5.1 Result analyzation

1. The accuracy of CNN is higher than the SVM model, which mean the error of predicton result of CNN is lower than the SVM model.
2. The precision of SVM is higher than CNN model, which mean the positive predictive value of SVM is higher than CNN model.
3. The recall of CNN model is higher than SVM model, which mean the sensitive of CNN model prediction is better than SVM model.
4. The F1-score of CNN model is higher the the SVM model, which mean the model score of CNN is higher than SVM model. The CNN model can predict more accurate than SVM model.
5. The running time of CNN is much lower than SVM model. That mean SVM is a very time consuming algorithm in sequence classification. Meanwhile, CNN will consume fewer time in training model. Thus, we can modify the model easily after prediction.

To summarize that CNN model is perform better than SVM in sequence classification.

Chapter 6

6. Conclusion

6.1 Achievement

In this report, I demonstrate the performance of deep learning model is better than support vector machine in sequence classification. Within the whole research I need to build up two model form zero. I need to search the DNA sequence data and convert the data form the human language to machine language. Then design the model structure of support vector machine and convolution neural network. Implement two model in jupyter notebook by programming. Train and test two model to get the result data. Finally can get a conclusion that the deep learning model can perform better the support vector machine model in sequence classification.

6.2 critical review

In the whole research, it let me know the machine learning is very close to our daily life. Machine learning not still are very far in application. In the beginning of the research, I tried to build the model at home. I used my not very power computer to train and test the model. It will consume near few day to get a trained model with small amount of data. However, when I borrow the high-performance computer with powerful GPU in my department. The training speed of model is increased few times. That let me know that one of the limitation of machine learning is memory and computing power. If the technical problem of the GPU or CPU is overcame and they become more powerful. The machine learning must have better

performance and extremely accuracy prediction. The machine learning maybe can popularize in our daily life.

Reference

[1]. Dl.acm.org. 2021. A brief survey on sequence classification | ACM SIGKDD Explorations Newsletter, pp. 41-45

[2]. Sundermeyer, M., Schluter, R., & Ney, H. (2012). LSTM Neural Networks for Language Modeling, pp194-197. Retrieved 14 October 2021,

[3]. S. V. M. Vishwanathan and M. Narasimha Murty, "SSVM: a simple SVM algorithm," Proceedings of the 2002 International Joint Conference on Neural Networks. IJCNN'02 (Cat. No.02CH37290), 2002, pp. 2393-2398 vol.3..

[4]. Yunqiang Chen, Xiang Sean Zhou and T. S. Huang, "One-class SVM for learning in image retrieval," Proceedings 2001 International Conference on Image Processing (Cat. No.01CH37205), 2001, pp. 34-37 vol.1

[5]. Peper, J. S., Brouwer, R. M., Boomsma, D. I., Kahn, R. S., & Hulshoff Pol, H. E. (2007). Genetic influences on human brain structure: a review of brain imaging studies in twins. Human brain mapping, 28(6), 464-473.

[6]. F. Q. Lauzon, "An introduction to deep learning," 2012 11th International Conference on Information Science, Signal Processing and their Applications (ISSPA), 2012, pp. 1438-1439

[7]. Bengio, Y., Goodfellow, I., & Courville, A. (2015). Deep Learning. Retrieved 14 October 2021, pp. 40-47

[8]. Hawkins, D. M. (2004). The problem of overfitting. Journal of chemical information and computer sciences, 44(1), 1-12.

[9]. Zhou, H., & Sun, Q. (2020, February). Research on Principle and Application of Convolutional Neural Networks. In IOP Conference Series: Earth and Environmental Science (Vol. 440, No. 4, p. 042055). IOP Publishing.

[10]. Wang, P., Chen, P., Yuan, Y., Liu, D., Huang, Z., Hou, X., & Cottrell, G. (2018, March). Understanding convolution for semantic segmentation. In 2018 IEEE winter conference on applications of computer vision (WACV) (pp. 1451-1460). IEEE.

[11]. Sun, M., Song, Z., Jiang, X., Pan, J., & Pang, Y. (2017). Learning pooling for convolutional neural network. Neurocomputing, 224, 96-104.

[12]. Jeczmionek, E., & Kowalski, P. A. (2021). Flattening Layer Pruning in Convolutional Neural Networks. Symmetry, 13(7), 1147.

[13]. Albawi, S., Mohammed, T. A., & Al-Zawi, S. (2017, August). Understanding of a convolutional neural network. In 2017 International Conference on Engineering and Technology (ICET) (pp. 1-6). Ieee.

[14]. T. O'Shea and J. Hoydis, "An Introduction to Deep Learning for the Physical Layer," in IEEE Transactions on Cognitive Communications and Networking, vol. 3, no. 4, pp. 563-575, Dec. 2017

[15]. Villon, S., Chaumont, M., Subsol, G., Villéger, S., Claverie, T., & Mouillot, D. (2016, October). Coral reef fish detection and recognition in underwater videos by supervised machine learning: Comparison between Deep Learning and HOG+ SVM methods. In International Conference on Advanced Concepts for Intelligent Vision Systems (pp. 160-171). Springer, Cham.

[16]. A. Shrestha and A. Mahmood, "Review of Deep Learning Algorithms and Architectures," in IEEE Access, vol. 7, pp. 53040-53065, 2019,

[17]. Liu, P., Choo, K. K. R., Wang, L., & Huang, F. (2017). SVM or deep learning? A comparative study on remote sensing image classification. Soft Computing, 21(23), 7053-7065.

[18]. Perlman, S. (2020). Another decade, another coronavirus.

[19]. Brutlag, D. L. (1980). Molecular arrangement and evolution of heterochromatic DNA. Annual review of genetics, 14(1), 121-144.

[20]. Raff, E. (2019). A step toward quantifying independently reproducible machine learning research. Advances in Neural Information Processing Systems, 32, 5485-5495.

[21]. Balachandran, P. V., Kowalski, B., Sehirlioglu, A., & Lookman, T. (2018). Experimental search for high-temperature ferroelectric perovskites guided by two-step machine learning. Nature communications, 9(1), 1-9.

[22]. Choong, A. C. H., & Lee, N. K. (2017, November). Evaluation of convolutionary neural networks modeling of DNA sequences using ordinal versus one-hot encoding method. In 2017 International Conference on Computer and Drone Applications (IConDA) (pp. 60-65). IEEE.

[23]. Lee, D., & Lee, J. (2007). Domain described support vector classifier for multi-classification problems. Pattern Recognition, 40(1), 41-51.

[24]. Yifan Peng, Anthony Rios, Ramakanth Kavuluru, Zhiyong Lu, Extracting chemical–protein relations with ensembles of SVM and deep learning models, Database, Volume 2018, 2018, bay073,

[25]. Deep Residual Learning for Image Recognition. Retrieved from http://arxiv.org/abs/1512.03385

[26]. He, F., Liu, T., & Tao, D. (2020). Why resnet works? residuals generalize. IEEE transactions on neural networks and learning systems, 31(12), 5349-5362.