EasyChair Preprint

№ 4499

Superhuman Intelligence: Mathematical Representation of a System As It Rises Through the Ranks of Intelligence.

Poondru Prithvinath Reddy

November 2, 2020

Superhuman Intelligence: Mathematical Representation of a System As It Rises Through the Ranks of Intelligence.

Poondru Prithvinath Reddy

ABSTRACT

Superhuman Intelligence (SI) is a stage of intelligence where machines will not only have superhuman strength and speed but also will have superhuman intelligence. SI is based on the idea that machines can imitate the human mind, their way of working to the extent that they can even supersede them. There are different paths of development for achieving Superhuman Intelligence and we have chosen mathematical approach from the first principles of the first seed AI system that will eventually grow up to be superintelligence and analyze how it will behave as it rises through the ranks of intelligence. In this paper, we implement object detection CNN-(R-SVM) combination where an AI algorithm (CNN) passes ‘Spatial Intelligence’ to fast machine learning (SVM) algorithm recursively which is a combination of ‘Spatial and Logical Intelligence’ for solving multiclass problems from large data sets that implements object detection for designing a better AI machine. The test results are encouraging with high accuracy and the model is therefore shown that a Spatial Intelligence is making a better AI when combined with intelligent vector algorithm recursively resulting in very high level of intelligence and if exposed to large data sets on a continuous basis then CNN-(R-SVM) can learn and develop cognitive abilities that will enable to grow up to be superintelligence.

INTRODUCTION

Superhuman Intelligence; An intellect that is much smarter than the best human brain in practically every field, including scientific creativity and general wisdom. Biological brains are unlikely to be the final stage of intelligence. Machines already have superhuman strength and speed – and one day they will have superhuman intelligence.

"**Superintelligence**" refers to the idea that steady advances in artificial intelligence, or machine (computer) intelligence, might one day result in creating a machine vastly superior to humans in reasoning and decision-making abilities.

Artificial Super Intelligence or ASI that has the capability to perform the tasks that are impossible for the human mind to think or do. It is that aspect of intelligence that is

more potent and refined than a human's intelligence and Superintelligence is capable of outperforming human intelligence. The human brain is made of neurons and is limited to some billion neurons. Superintelligence, therefore challenges this trait, which knows no limit.

There are different types of intelligence and we look at particular two types of intelligence mainly Spatial Intelligence and Logical-Mathematical Intelligence which are as below:

- Logical-Mathematical. Interested in patterns, categories and relationships.
- Spatial. Think in images and pictures.

Logical-Mathematical Intelligence: capacity to use numbers in an effective way, and to transform with dexterity, different reasoning chains. A person with a good development of logical-mathematical intelligence, highlights in the resolution of problems, to carry out complicated mathematics calculations and logical reasoning, as scientists, engineers, economists and others.

Spatial Intelligence: capacity to perceive the visual and spatial world, and to transform or to recognize its elements. This intelligence includes several informal abilities, such as: ability to use the imagination and then transform it, ability to visualize colors, lines, shapes and figures and others, ability to produce graphic likeness of spatial information and the orientation ability. Some professionals that possess this intelligence in a high level of development are the designers, architects, mathematicians and engineers.

Support Vector Machine (SVM)

Linear SVM is the newest extremely fast machine learning (data mining) algorithm for solving multiclass classification problems from ultra large data sets that implements an original proprietary version of a cutting plane algorithm for designing a **linear** support vector machine.

In machine learning, **Support Vector Machine** (**SVM**) is a **non**-probabilistic, **linear**, binary classifier used for classifying data by learning a hyperplane separating the data. Classifying a **non**-**linearly** separable dataset using a **SVM** – a linear classifier: However, it can be used for classifying a non-linear dataset also.

Linear classifier (SVM) is used when number of features are very high, e.g., document **classification**. This is because **Linear** SVM gives almost similar accuracy as **non linear** SVM but **Linear** SVM is very very fast in such cases and **non**-**linear classifier** is useful when data is not linearly separable.

METHODOLOGY

Superintelligence

However, for AI to lead to superintelligence, we don't necessarily need to develop superintelligent AI ourselves and we only need to develop an AI that is marginally better at developing AI's than we are.

When that happens, we have an AI that makes better AI's that make better AI's — all in the blink of a human eye and when AI's master this, superintelligence might be just around the corner.

Support Vector Machine (SVM)

The support vector machine (SVM) was developed by for binary classification. Its objective is to find the optimal hyperplane $f(w, x)=w \cdot x + b$ to separate two classes in a given dataset, with features $x \in R^m$.

SVM learns the parameters w by solving an optimization problem (Eq. 1).

$$min \frac{1}{p} w^T w + C \sum_{i=1}^{p} \max(0,1 - y_i' (w^T x_i + b)) \quad (1)$$

Where $w^T w$ is the Manhattan norm (also known as L1 norm), C is the penalty parameter (may be an arbitrary value or a selected value using hyper-parameter tuning), y' is the actual label, and $w^T x + b$ is the predictor function. Eq. 1 is known as L1-SVM, with the standard hinge loss. Its differentiable counterpart, L2-SVM (Eq.2), provides more stable results.

$$min \frac{1}{p} \|w\|_2^2 + C \sum_{i=1}^{p} \max(0,1 - y_i' (w^T x_i + b))^2 \quad (2)$$

Where $\|w\|2$ is the Euclidean norm (also known as L2 norm), with the squared hinge loss.

Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) is a class of deep feed-forward artificial neural networks which is commonly used in computer vision problems such as image classification. The distinction of CNN from a "plain" multilayer perceptron (MLP) network is its usage of convolutional layers, pooling, and non-linearities such as tanh, siдmoid, and ReLU.

The convolutional layer (denoted by CONV) consists of a filter, for instance,5×5×1(5 pixels for width and height, and 1 because the images are in grayscale). Intuitively speaking, the CONV layer is used to "slide" through the width and height of an input image, and compute the dot product of the input's region and the weight learning

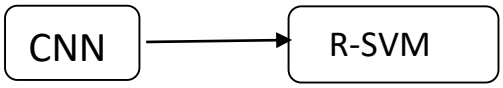
parameters. This in turn will produce a 2-dimensional activation map that consists of responses of the filter at given regions. Consequently, the pooling layer (denoted by POOL) reduces the size of input images as per the results of a CONV filter. As a result, the number of parameters within the model is also reduced – called down-sampling

Lastly, an activation function is used for introducing non-linearities in the computation. Without such, the model will only learn linear mappings. The commonly-used activation function is the ReLU. ReLU is commonly-used over tanh and sigmoid for it was found out that it greatly accelerates the convergence of stochastic gradient descent compared the other two functions. Furthermore, compared to the extensive computation required by tanh and sigmoid, ReLU is implemented by simply thresholding matrix values at zero (see below Eq. 3).

$$f(h_\Theta(x)) = h_\Theta(x)^+ = \max(0, h_\Theta(x)) \qquad (3)$$

ARCHITECTURE

We define following structure of CNN-(R-SVM) model for applying spatial-logical intelligence with recursive SVM implementations to achieve optimal method of object localization and recognition.



Recursive CNN-SVM Architecture for Image Classification

MNIST is an established standard hand written digit classification dataset that is widely used for benchmarking deep learning models. However, we have used the Fashion-MNIST dataset. The said dataset consists of images having the same distribution, the same number of classes, and the same color profile as MNIST. Also, it is having 10,000 training examples and 10,000 test cases.

The Deep Artificial Neural Network

We used two convolutional layers:

- The first layer will have 32-5 x 5 filters,
- The second layer will have 64-5 x 5 filters

In addition, there are two max-pooling layers each of size 2 x 2.

We used a RELU as our activation function which simply takes the output of max_pool and applies RELU.

Fully connected layer:

Just like any other layer, we declare weights and biases as random normal distributions. In fully connected layer, we take all the inputs, do the standard operation on it. The Fully Connected Layer has 1024 Hidden Neurons.

We added Dropout into the network to overcome the problem of overfitting to some extent and also to improve the training and validation accuracy.

The final layer is Output Layer with 10 Output Classes.

At the last layer of the CNN, instead of the conventional softmax function with the cross entropy function (for computing loss), the L2-SVM is implemented. That is, the output shall be translated to the following case $y \in \{-1, +1\}$, and the loss is computed by Eq. 2.The weight parameters are then learned using Adam.

Machine Learning methods for feature selection and classification have been playing active roles in analyzing high-throughput data. We used both normal linear SVM and recursive support vector machine (R-SVM) to select input features for classification. The proposed R-SVM algorithm will recursively classify the training samples with SVM and select features according to their weights in the SVM classifier.

TEST RESULTS

Image Classification is about classifying objects in an image and the test results show good accuracy between training and validation data. However the CNN algorithm needs a lot of regions to predict accurately and hence high computation time.

We compared two methods i.e. SVM and R-SVM, the R-SVM adopting recursive procedures to select features in SVM classifiers. Although the two methods (SVM and R-SVM) did not differ significantly in their validation performances, it appears that R-SVM is more robust and can recover more informative features. The proposed R-SVM method is suitable for analyzing high-throughput data and it outperforms SVM in the robustness and in the ability to recover informative features.

CONCLUSION

Superhuman Intelligence (SI) is a stage of intelligence where machines will not only have Superhuman strength and speed but also will have superhuman intelligence. In order to achieve this, SI will have to be more intelligent by improving intelligent abilities of the artificial machines. In this paper, we have implemented object detection CNN-(R-SVM) architecture where a lesser machine with Spatial Intelligence pass intelligence to high level machine learning algorithm combined with Spatial and Logical Intelligence to make better AI recursively. The test results show that it is possible to build artificial machines that have high-level intelligence by combining different methods of intelligence logically which when exposed to large data sets on a continuous basis may grow up to be superintelligence.

REFERENCE

1. https://github.com/
2. Abien Fred M. Agarap “An Architecture Combining Convolutional Neural Network(CNN) and Support Vector Machine (SVM) for ImageClassification”. https://arxiv.org/pdf/1712.03541.pdf