Project Report for ESI6247

Statistical Design Models

HYPERPARAMETER OPTIMIZATION USING EXPERIMENTAL DESIGN TO IMPROVE THE PERFORMANCE IN IMAGE CLASSIFICATION

Anusha Durgam (U93981106)

Master's in Data Intelligence

Department of Industrial and Management Systems Engineering

College of Engineering, University of South Florida

Dr. Tapas K. Das

December 5, 2024

CONTENTS

Executive Summary

Deep learning relies heavily on neural networks (NNs), which are modelled after the human brain. Convolutional neural networks (CNNs) excel at image classification and object recognition. However, optimal performance requires precise tuning of hyperparameters like learning rate, batch size, and architectural configurations.

The purpose of this project is to improve validation accuracy by optimizing the hyperparameters of a convolutional neural network (CNN) using Design of Experiments (DOE). The study is carried out individually, beginning with a baseline experiment to establish initial performance metrics. Key hyperparameters, such as the kernel (filter) size in Conv2D layers, were systematically varied, and validation accuracies were calculated. Graphical analysis was used to identify the low and high hyperparameter values.
To efficiently test combinations of these levels, we used a fractional factorial design. The obtained data was used to develop a regression model that predicted validation accuracy. The model's predictions were then compared to actual CNN performance, revealing discrepancies and providing insights into hyperparameter significance and model assumptions.

As a result, the optimized model had a R^2 accuracy of 99.83%, The study emphasized the importance of precise hyperparameter tuning in optimizing model efficiency. This project demonstrates DOE's effectiveness in neural network optimization by providing a structured, scalable framework for hyperparameter tuning in deep learning applications.

INTRODUCTION

Neural networks, particularly deep learning models, have emerged as a foundation of modern artificial intelligence (AI), driving advances in decision-making across a wide range of fields. These models are intended to mimic the structure of the human brain, with layers of interconnected nodes (neurons) processing and learning from massive amounts of data. Neural networks allow systems to make highly accurate predictions and decisions for every scenario by extracting complex patterns from large datasets (LeCun, Bengio, & Hinton, 2015). Deep learning, a subset of machine learning, uses multi-layered networks known as deep neural networks (DNNs) to improve performance by learning hierarchical feature representations from data (Goodfellow, Bengio, & Courville, 2016).

Convolutional Neural Networks (CNNs) are especially effective at processing structured data like images and videos. CNNs are intended to automatically detect patterns such as edges, textures, and shapes at various levels of abstraction, making them particularly useful for image classification, object detection, and facial recognition (Krizhevsky, Sutskever, & Hinton 2012). This ability to automatically learn relevant features from raw data has resulted in significant improvements in a variety of domains, including healthcare (Esteva et al., 2019) and autonomous vehicles (Bojarski et al., 2016).

Despite their power, neural network’s success is heavily dependent on careful tuning of hyperparameters such as learning rate, batch size, and architecture-related parameters (Bengio, 2012). Inaccurate hyperparameter settings can significantly degrade the model's performance, resulting in poor generalization and prediction accuracy. As a result, optimizing these hyperparameters is an important step in the model development process, especially when working with large-scale datasets. This project aims to optimize hyperparameters for Convolutional Neural Networks (CNNs) using the CIFAR-10 dataset, a well-known image classification benchmark. This study uses Design of Experiments (DOE) methods to identify the most significant hyperparameters that affect CNN performance and investigate their interactions to improve the model's classification accuracy.

PROBLEM DESCRIPTION

This project focuses on improving the accuracy of image classification by fine-tuning essential hyperparameters, utilizing design of experiments to systematically optimize them.

The dataset we will be using is CIFAR-10 dataset, categorized into 10 distinct classes, providing a benchmark for image classification tasks.

1. Ship
2. Truck
3. Bird
4. Frog
5. Automobile
6. Deer
7. Airplane
8. Horse
9. Dog
10. Cat

Table 1: Dataset categories

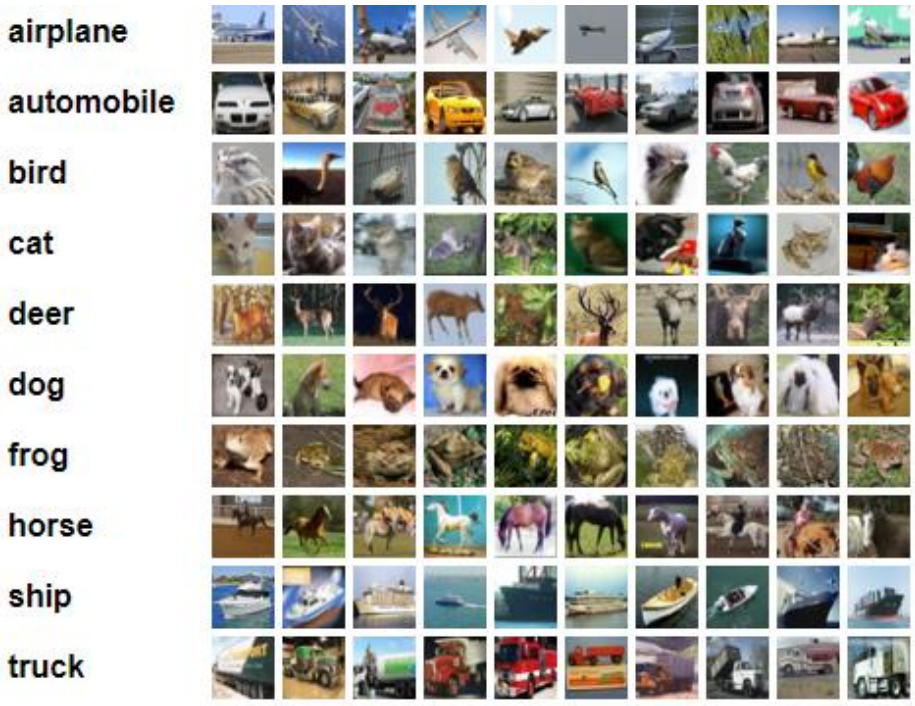


Figure 1: CIFRA-10 DATASET

The code provided involves 8 hyperparameters that influence the performance of the neural network. By adjusting these parameters, we aim to optimize the model's accuracy for better image classification. Those 8 parameters are mentioned below:

- Kernel Size
- Number of Convolutional Layers
- Number of Filter Conv 2D
- Number of Neurons
- Learning Rate
- Number of Epochs
- Batch size &
- Momentum

Table 2: Hyperparameters

The code provides with the base values for this hyperparameters which are mentioned below:

```
kernel_size_Conv2D = 11
number_of_convolutional_layers = 1
number_of_filter_conv2D = 64
activation = 'sigmoid'
number_of_neurons = 300
learning_rate = 0.001
number_of_epochs = 10
batch_size = 10
momentum = 0.8
```

Figure 2: Base values for the Hyperparameters

PROBLEM OBJECTIVE

The goal of this project is to improve image classification accuracy by fine-tuning key hyperparameters that affect the performance of a Convolutional Neural Network (CNN). We intend to identify the most significant factors by designing an experiment based on a half-fractional factorial model. By optimizing these hyperparameters, we will create a model that accurately predicts outcomes, which will then be validated using real neural network data. To evaluate the improvements made, we will compare the performance of our optimized model to that of the original CNN.

METHODOLOGY

The methodology for this project is designed to optimize the hyperparameters of a Convolutional Neural Network (CNN) using the CIFAR-10 dataset. The process involves systematic experimentation to explore the impact of various hyperparameter values on model performance.

1. Selection of Hyperparameters

The following eight hyperparameters are designated for analytical purposes, with each having a baseline value and an allowable range for experimentation. These parameters are crucial for tuning the performance of the CNN, and we will vary them systematically to identify their impact on the model's accuracy:

Hyperparameter	Letter	Baseline Value	Allowable Range
Kernel Size	A	11	1, 3, 5, 7, 11 (integer)
Number of Convolutional Layers	B	1	1-4 (integer)
Number of Filters	C	64	16, 32, 64, 128, 256 (integer)
Number of Neurons	D	300	50-300 (integer)
Learning Rate	E	0.001	0.0001 - 1 (continuous)
Number of Epochs	F	10	1-40
Batch Size	G	10	1-100 (integer)
Momentum	H	0.8	0.5 - 0.99 (continuous)

Table 3: Hyperparameters Range

2. Hyperparameter Experimentation

Each hyperparameter will be varied within its defined range, while the others will remain at their baseline values. For each combination of hyperparameters, we will track the CNN model's accuracy. A graph will be plotted for each hyperparameter to visualize how different values correlate with the model's performance. This analysis will allow us to pinpoint the most significant factors that affect accuracy.

3. Factorial Experiment Design

To determine the interactions and significance of the hyperparameters, we will use a **Half-Fractional Factorial Design**. This method enables us to explore the influence of all the factors in a limited number of experiments by reducing the number of combinations while still capturing key interactions.

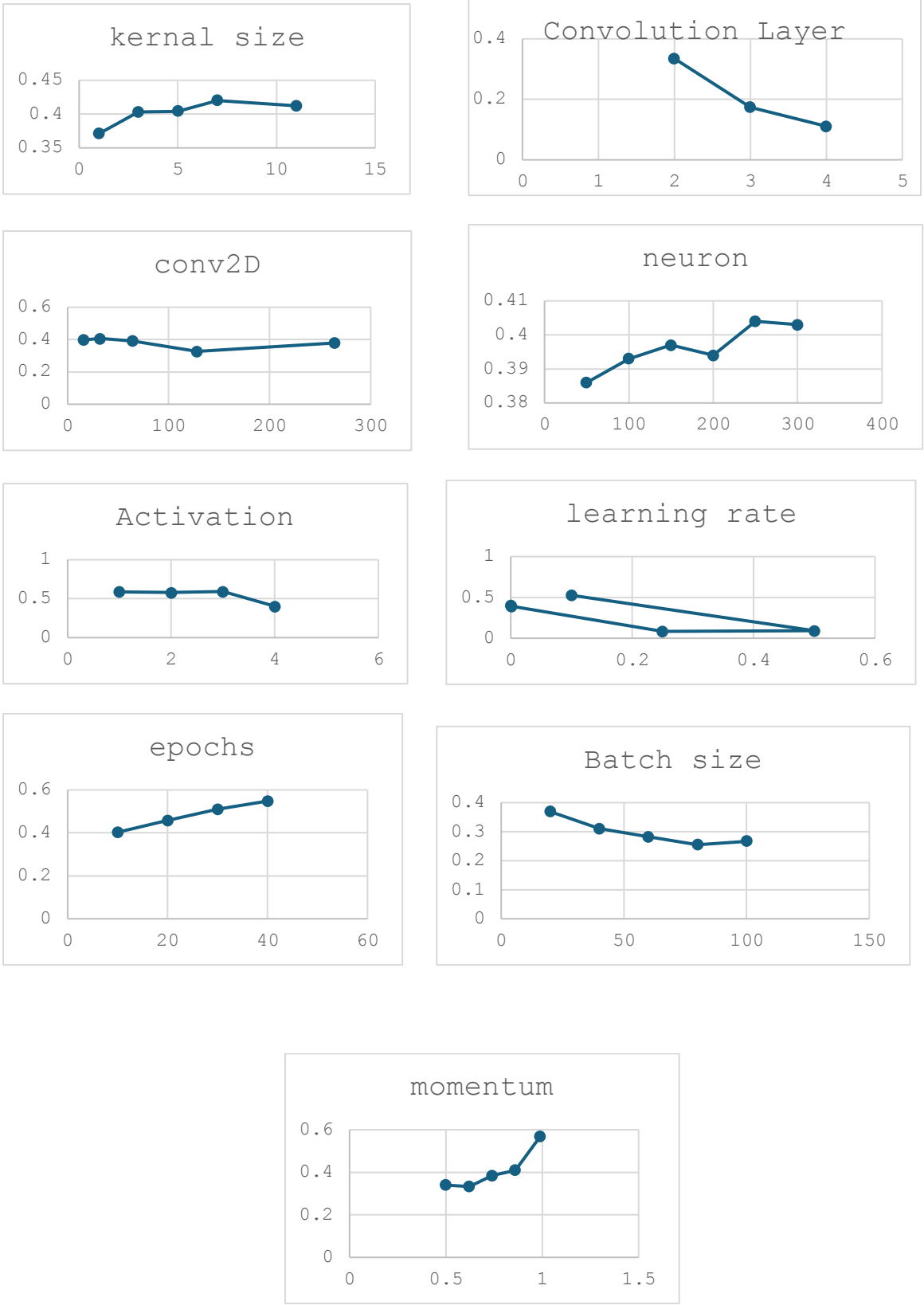
Steps involved in this design include:

- **Designing the Fractional Factorial Experiment**: We will perform a half-fractional factorial experiment by reducing the total number of combinations, enabling us to focus on the most significant interactions.
- **ANOVA Analysis**: We will conduct an Analysis of Variance (ANOVA) to identify which factors and their interactions have the greatest influence on model accuracy.
- **Regression Model Development**: A regression equation will be formulated to describe the relationship between the hyperparameters and the model's accuracy. The R^2 value will be computed to assess the model's effectiveness.
- **Validation**: The model is validated by comparing the validation accuracy obtained from the regression model with the actual accuracy from the neural network using the same hyperparameter combinations.

EXPERIMENT AND DATA COLLECTION

1.Finding Accuracies

In the experiment, we evaluated the effect of individual hyperparameters on model accuracy by varying one hyperparameter at a time within its defined range and observing the resulting accuracy. This process allowed us to identify the high and low values for each hyperparameter, which are used in the subsequent fractional factorial trials. The following graphs were plotted to show the relationship between each hyperparameter and the accuracy.



Figures 4: Hyperparameters vs Corresponding accuracy Graphs

Based on the plots we can find the low and values for each of the hyperparameters. The above are the values obtained after plotting:

Parameters	LOW	HIGH
Kernel Size	1	7
Number of Convolutional Layers	4	1
Number of Filters	128	32
Number of Neurons	50	250
Learning Rate	0.25	0.1
Number of Epochs	10	40
Batch Size	80	20
Momentum	0.622	0.99

Table 4: High and Low values of hyperparameters

2.Half Fractional Factorial Experiment

This table below outlines the five hyperparameters randomly chosen for the experiment. These factors were used to perform the 2^5 full factorial design and subsequently reduced to a 1/2 fractional factorial experiment (2^5 - 1) for more efficient analysis.

Hyperparameter	Factor	Low value	High value
A	Kernal size	1	7
B	filter conv 2D	128	32
C	learning rate	0.25	0.1
D	batch size	80	20
E	momentum	0.622	0.99

Table 5: Factors selected

After constructing the **2^5 - 1** fractional factorial design, we proceeded to build the **ANOVA (Analysis of Variance) table** to evaluate the significance of each factor and its interactions.

Using the data obtained from the experiments, we calculated the **regression equation** to model the relationship between the hyperparameters and the accuracy. Finally, we computed the **R^2 value** to assess the goodness of fit and determine how well the model explained the variation in accuracy.

3. Assumption and Validation

For validation, we compared the predicted accuracy from the regression model with the actual results obtained from the neural network, assessing the model's effectiveness in real-world scenarios.

DATA ANALYSIS

Half-Fractional Factorial Experiment

The data analysis consisted of calculating a regression equation based on the half-fractional factorial experiment results. The ANOVA table was used to determine the significant hyperparameters influencing the CNN model's accuracy. With f0>fcritical for all factors that we were calculating, are found that the hyperparameters had a significant effect on performance. The regression equation from the experiment achieved a R^2 value of 0.99833, indicating a model accuracy of 99.83%.

The significant factors are: A ,B,C,D,E,AB,AC,AD,AE,BC,BD,BE,CD,CE,DE.

R^2= 0.998332

Regression equation:

Y=0.245313+(0.0565/2)*X1+(0.1195/2)*X2+(0.0865/2)*X3+(0.0475/2)*X4+(-0.29313/2)*X5+(-0.01213/2)*X1X2+(0.020375/2)*X1X3+(-0.08263/2)*X2X3+(0.077375/2)*X1X4+(-0.01338/2)*X2X4+(0.014375/2)*X3X4+(-0.047/2)*X1X5+(-0.12675/2)*X2X5+(-0.0745/2)*X3X5+(-0.04325/2)*X4X5

Assumption and Validation

Coded the assumption values using code **(Xcode(H+L)/2)/((H−L)/2** between the actual hyperparameter's high and low values. The accuracy of five assumed values was verified. After receiving the coded beta values, we created the regression equation for each row using the same factors and interactions as in the one-half factorial experiment. When comparing code and assumption accuracy, the third assumption was discovered to be the best model (17.32%).

We can consider this the best model because the smaller the difference between coded accuracy and assumption accuracy, the higher the chance of the model being the best.

coded accuracy	Assumption Accuracy	absolute accuracy	(coded/assumption)%
0.314238951	0.484	0.350745969	35.07459689
0.132897616	0.108	0.230533486	23.0533486
0.6382505	0.544	0.173254596	17.32545955

Table 6: Assumption and validation accuracy comparison

RESULTS

The R-squared value of 99.83% obtained in our model indicates that the system detects the image with high accuracy. This extremely high R-squared value demonstrates that the model accurately captures the relationships between the variables and their effects on the outcome. The model's performance across the specified factor settings adds to the reliability of the experimental results. After validating the model with various hyperparameter combinations and the regression model, we discovered that the third assumption provided the highest accuracy. As a result, we determined that the third model was the best.

CONCLUDING REMARKS

The fractional factorial experiment provided useful information about the factors influencing our model's performance. Using Design of Experiments (DOE) allowed us to systematically identify the most important variables and their intricate interactions, guiding us to optimize the model for maximum efficiency. The experiment emphasized the importance of careful hyperparameter selection and validation, demonstrating how even minor changes can have a significant impact on results. This process demonstrates the effectiveness of DOE in refining machine learning models for real-world applications.