Progress Journal for CAMS Project, Team Radiant

Sara Haman, Kalani Stanton, Vivienne Prince

2021-11-25

Contents

Preface

This project was completed in accordance with the requirements for the Practical Data Science course in the Data Science Masters program at New College of Florida.

Chapter 1

Introduction

This is indeed a book. Wow.

Chapter 2

Documentation

such an absolute book, wow.

2.1 Getting Started

2.1.1 Adding a Model

If you would like to experiment with your own architecture variations

2.1.2 Training A Model

2.1.3 Directory Structure

```
|_README.md
|_train_experiment.py
|_train_LatLng.py
|
|_Data
| |_
|_Models
```

2.1.3.1 Data

2.1.3.2 Results

2.2 Updates and Modifications

Chapter 3

Model Architectures

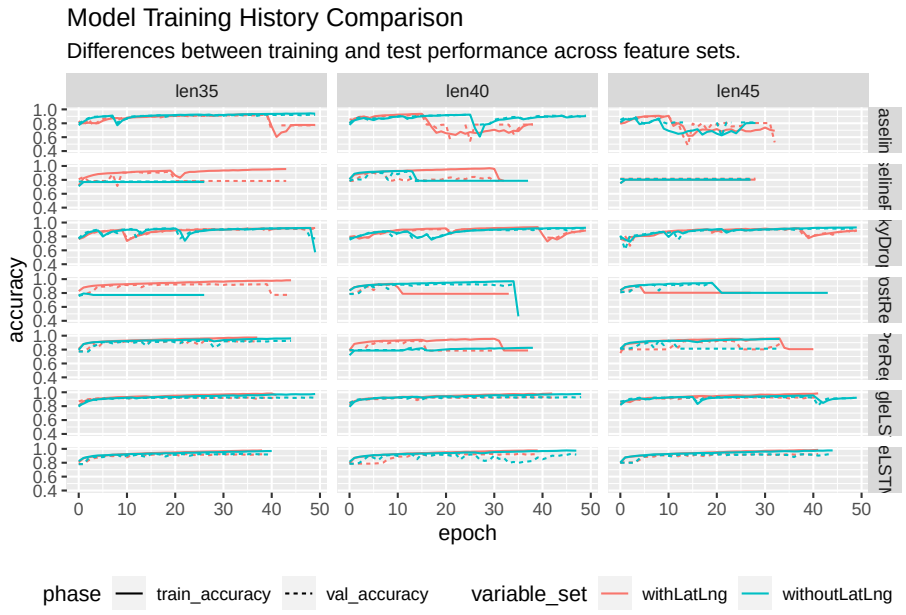
```
library(tidyverse)
```

```
featureset_df = read_csv('data/featuresets_ModelHistories.csv')

accuracy_df <- featureset_df %>%
  rename(train_loss = loss, train_accuracy = accuracy) %>%
  select(-c("val_loss", "train_loss")) %>%
  pivot_longer(cols = c(train_accuracy, val_accuracy), names_to = "phase", values_to = "accuracy"

loss_df <- featureset_df %>%
  rename(train_loss = loss, train_accuracy = accuracy) %>%
  select(-c("val_accuracy", "train_accuracy")) %>%
  pivot_longer(cols = c(train_loss, val_loss), names_to = "phase", values_to = "loss")
```

```
ggplot(accuracy_df, aes(x = epoch, y = accuracy, color = variable_set, linetype = phase)) +
  geom_line(linewidth = 10) +
  ylim(0.4,1.0) +
  facet_grid(vars(model), vars(len_instance)) +
  ggtitle("Model Training History Comparison", subtitle = "Differences between training and test
```


Model Training History Comparison
Differences between training and test performance across feature sets.
len35
len40
len45
accuracy
epoch
phase
train_accuracy
val_accuracy
variable_set
withLatLng
withoutLatLng

Chapter 4

Model Evaluation

4.1 Feature set data

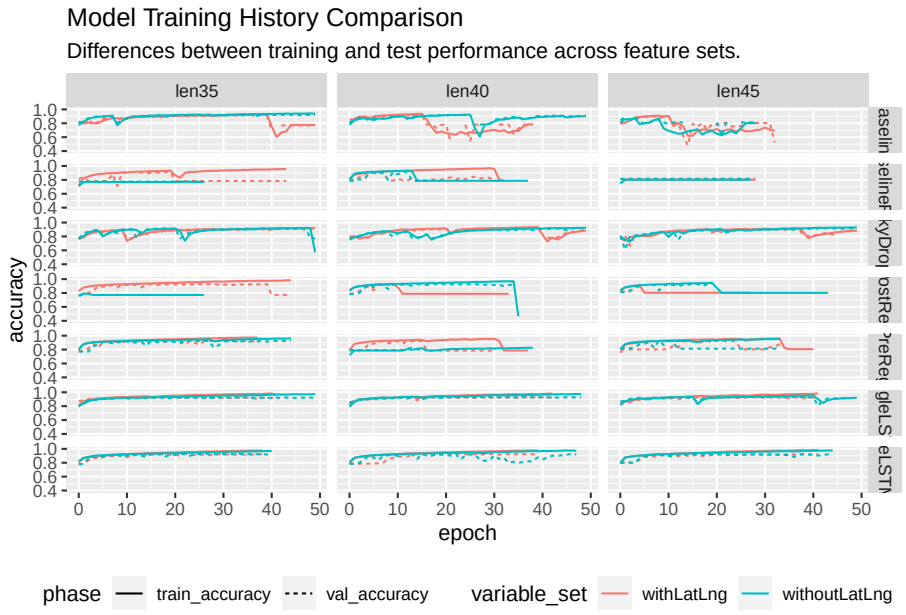
```
library(tidyverse)
```

```
featureset_df = read_csv('data/featuresets_ModelHistories.csv')

accuracy_df <- featureset_df %>%
  rename(train_loss = loss, train_accuracy = accuracy) %>%
  select(-c("val_loss", "train_loss")) %>%
  pivot_longer(cols = c(train_accuracy, val_accuracy), names_to = "phase", values_to = "accuracy"

loss_df <- featureset_df %>%
  rename(train_loss = loss, train_accuracy = accuracy) %>%
  select(-c("val_accuracy", "train_accuracy")) %>%
  pivot_longer(cols = c(train_loss, val_loss), names_to = "phase", values_to = "loss")
```

```
ggplot(accuracy_df, aes(x = epoch, y = accuracy, color = variable_set, linetype = phase)) +
  geom_line(linewidth = 10) +
  ylim(0.4,1.0) +
  facet_grid(vars(model), vars(len_instance)) +
  ggtitle("Model Training History Comparison", subtitle = "Differences between training and test
```


Model Training History Comparison
Differences between training and test performance across feature sets.
len35
len40
len45
accuracy
epoch
phase
train_accuracy
val_accuracy
variable_set
withLatLng
withoutLatLng