Progress Journal for CAMS Project, Team Radiant

Sara Haman, Kalani Stanton, Vivienne Prince

2021-11-24

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

Chapter 2

Progress Reports

```
library(plyr)
library(tidyverse)
library(Hmisc)
library(httr)
library(jsonlite)
library(tidyverse)
#install.packages("tmap")
library(tmap)
library(leaflet)
library(treemap)
library(kableExtra)

`%notin%` <- Negate(`%in%`)

MakePretty <- function(x) {
  x %>% kbl(align = "c") %>% kable_material()
}
```

2.1   Week 1 (8/30 - 9/4)

some stuff

2.1.1 Achievements

2.1.2 Aggravations

Error from inside the model, possibly due to version incompatibilities. May have to force inputs to numpy array.

### 2.1.3 Aims

2.1.4 Attitude

2.2 Week 2

Trained the model on intervals 30 - 50

```r
LenInstance <- c(30,35,40,45,50, 55)
maxAcc <- c(0.9238, 0.9331, 0.9396, 0.9417,0.8177, 0.9097)
maxValAcc <- c(0.9164, 0.9349, 0.8876, 0.9168, 0.8240, 0.9107)
epoch <-  c(13, 12, 15, 49, 9, 17)

modelResults <- data.frame(cbind(LenInstance,maxAcc, maxValAcc, epoch))

modelResults <- pivot_longer(modelResults, cols = c(maxAcc, maxValAcc))
```

```r
library(ggplot2)

ggplot(modelResults, aes(x = epoch, y = value, color = factor(LenInstance), shape = na
```
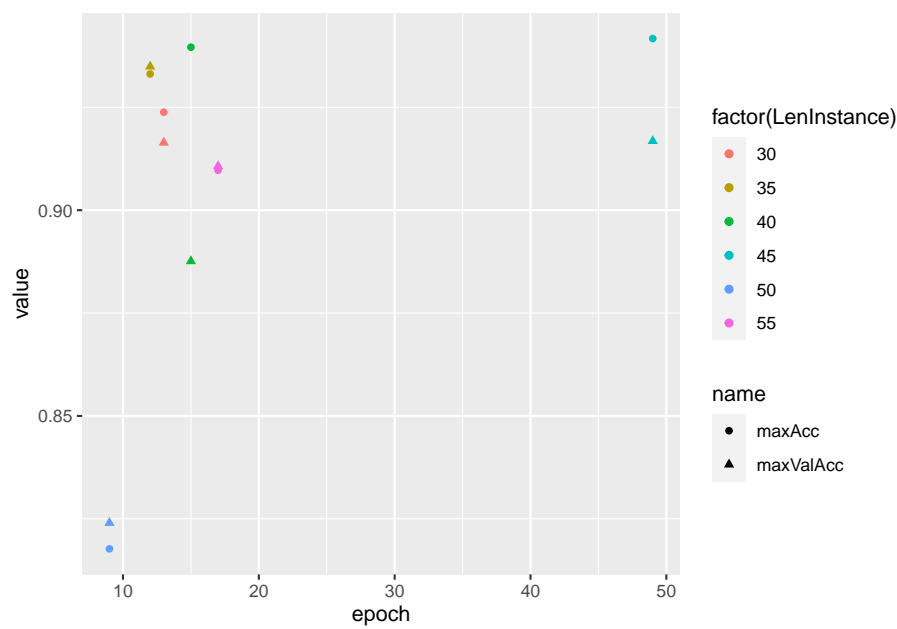
Chapter 3

Model Performance

```r
library(tidyverse)
```

```r
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")
```

```r
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.