TA-Q3

Jakki Seshapanpu (11915005),

Manoj Sharma (11915050),

Utkarsh Agrawal (11915020)

**ui.R**

#

#

# Jakki Seshapanpu (11915005)

# Manoj Sharma (11915050)

# Utkarsh Agrawal (11915020)

#

#

library("shiny")

require(dplyr)

require(magrittr)

require(tidytext)

library(udpipe)

library(textrank)

library(lattice)

library(igraph)

library(ggraph)

library(ggplot2)

library(wordcloud)

library(stringr)

library(topicmodels)

library(dplyr)

library(RColorBrewer)

library(stringr)

# Define ui function

shinyUI(

fluidPage(

titlePanel("Building a Shiny App around the UDPipe NLP workflow"),

sidebarLayout(

sidebarPanel(

fileInput("file", "Upload a text file "),

selectInput("select\_var", "Model",

c("english-ewt", "english-gum", "english-lines", "english-partut"), selected = "english-ewt"),

checkboxGroupInput("input\_selection", "select list of Universal part-of-speech tags (upos)",

c("adjective (ADJ)" = "adj",

"noun(NOUN)" = "noun",

"proper noun (PROPN)" = "propn",

"adverb (ADV)" = "adv",

"verb (VERB)" = "verb"),

selected = c("adjective (ADJ)" = "adj",

"noun(NOUN)" = "noun",

"proper noun (PROPN)" = "propn"),

inline = FALSE)),

mainPanel(

tabsetPanel(type = "tabs",

tabPanel("Description about this app", h4(p("Workflow about using this app")),

p("This Shiny App related to UDPipe NLP workflow.\n\n

Upload the file by clicking the browse button on the left side. \n\n

After uploading the file, select the model. \n\n

Then select the Universal part-of-speech tags from the selection check boxes. \n\n

Now select on tab(annotated documents or wordclouds or co-ocurrences) to see the results ", align = "justify")),

tabPanel("Annotated documents",

h4(p("Download Annotated document file")),

downloadButton('downloadData1', 'Download Annotated document file'),br(),br(),

h4(p("Display of the annotated documents")),

dataTableOutput('annotated\_df')),

tabPanel("Word clouds",

h4("Nouns Word Cloud"),

plotOutput("nouns\_wordcloud",height = 700, width = 700),

h4("Verbs Word Cloud"),

plotOutput("verbs\_wordcloud",height = 700, width = 700)),

tabPanel("Co-occurrences ",

h4("Co occurrences Graph "),

plotOutput('Co\_occurrences\_Graph'))

) # end of tabsetPanel

)# end of main panel

) # end of sidebarLayout

) # end if fluidPage

) # end of UI

Server.R

#

#

# Jakki Seshapanpu (11915005)

# Manoj Sharma (11915050)

# Utkarsh Agrawal (11915020)

#

#

library(shiny)

shinyServer(function(input, output) {

###########################################################

dataset <- reactive({

if (is.null(input$file)) {return(NULL)}

else {

dataset = readLines(input$file$datapath)

dataset = str\_replace\_all(dataset, "<.\*?>", "")

return(dataset)}

})

###########################################################

## annotated df

annotated\_df\_calculator <- reactive({

m\_type = udpipe\_load\_model("D:\\2\_p\\Jakki\_Personnel\\Studies\\CBA\\T1\\TA\\GroupProject\\TA-Q3\\Q3\\english-ewt-ud-2.4-190531.udpipe")

x <- udpipe\_annotate(m\_type, x = dataset())

x <- select(as.data.frame(x),-sentence)

x

head(x,100)

return(x)

})

output$annotated\_df <- renderDataTable({ annotated\_df\_calculator() })

###########################################################

#wc(nouns)

nouns\_wordcloud\_calculator <- reactive({

df = annotated\_df\_calculator()

all\_nouns = df %>% subset(., upos %in% "NOUN"); all\_nouns$token[1:20]

top\_nouns = txt\_freq(all\_nouns$lemma)

wordcloud(words = top\_nouns$key,

freq = top\_nouns$freq,

min.freq = 1,

max.words = 100000,

random.order = FALSE,

colors = brewer.pal(6, "Dark2"))

}) #end of function

output$nouns\_wordcloud<-renderPlot({ nouns\_wordcloud\_calculator() })

###########################################################

#wc(verbs)

verbs\_wordcloud\_calculator <- reactive({

df = annotated\_df\_calculator()

all\_verbs = df %>% subset(., upos %in% "VERB"); all\_verbs$token[1:20]

top\_verbs = txt\_freq(all\_verbs$lemma)

wordcloud(words = top\_verbs$key,

freq = top\_verbs$freq,

min.freq = 1,

max.words = 100000,

random.order = FALSE,

colors = brewer.pal(6, "Dark2"))

}) #end of function

output$verbs\_wordcloud<-renderPlot({ verbs\_wordcloud\_calculator() })

###########################################################

#co occurances graph

Co\_occurrences\_Graph\_calculator <- reactive({

x = annotated\_df\_calculator()

text\_cooc <- cooccurrence(

x = subset(x, upos %in% c("NOUN", "ADJ")),

term = "lemma",

group = c("doc\_id", "paragraph\_id", "sentence\_id"))

wordnetwork <- head(text\_cooc, 50)

wordnetwork <- igraph::graph\_from\_data\_frame(wordnetwork)

ggraph(wordnetwork, layout = "fr") +

geom\_edge\_link(aes(width = cooc, edge\_alpha = cooc), edge\_colour = "orange") +

geom\_node\_text(aes(label = name), col = "darkgreen", size = 4) +

theme\_graph(base\_family = "Arial Narrow") +

theme(legend.position = "none") +

labs(title = "Cooccurrences Graph")

}) #end of function

output$Co\_occurrences\_Graph<-renderPlot({ Co\_occurrences\_Graph\_calculator() })

###########################################################

#download annotated document file

output$downloadData1 <- downloadHandler(

filename = function() { "AnnotatedDocument.csv" },

content = function(file) {

write.csv(readLines("data/AnnotatedDocument.csv"), file)

}

)

})