#下载包

install.packages("ggplot2")

install.packages("scales", dependencies = FALSE)

install.packages("GGally", dependencies = FALSE)

install.packages("plyr", dependencies = FALSE)

#导入包

library(ggplot2)

library(scales)

library(grid)

#导入数据

data <- read.csv("D://desktop/WA\_Fn-UseC\_-Telco-Customer-Churn.csv",header = T)

attach(data)

class(data)

#查看数据

dim(data)

head(data)

str(data)

#查看缺失值

is.null(data)

#计算缺失值个数占比

na<-is.na(data$TotalCharges)

na[na==FALSE]<-0

na[na==TRUE]<-1

na.rate<-as.numeric(sum(na)/length(na))

na.rate

#剔除缺失值

data<-na.omit(data)

str(data)#查看数据类型

data$SeniorCitizen[data$SeniorCitizen==0]<-"No"

data$SeniorCitizen[data$SeniorCitizen==1]<-"Yes"

#转换为因子变量

for(i in c(2:5,7:18,21)) {

data[i]<-as.factor(unlist(data[i]))

}

str(data)

for(i in 10:15){

print(xtabs(~ Churn + get(names(data)[i]), data = data))

}

# 将"No internetserive"并入"No"这一属性值

levels(data$OnlineSecurity)[2] <- "No"

levels(data$OnlineBackup)[2] <- "No"

levels(data$DeviceProtection)[2] <- "No"

levels(data$TechSupport)[2] <- "No"

levels(data$StreamingTV)[2] <- "No"

levels(data$StreamingMovies)[2] <- "No"

options(digits=4)

#饼图

ggplot(data, aes(x = "" ,fill = Churn))+

geom\_bar(stat = "count", width = 0.5, position = 'stack')+

coord\_polar(theta = "y", start=0)+

geom\_text(stat="count",

aes(label = scales::percent(after\_stat(count), 0.01)),

size=4, position=position\_stack(vjust = 0.5)) +

theme(

panel.background = element\_blank(),

axis.title = element\_blank(),

axis.text = element\_blank(),

axis.ticks = element\_blank()

)

#条形图

ggplot(data , aes(x =Churn , y =after\_stat(count) ,fill=Churn)) +

geom\_bar(stat = "count", width = 0.5, position = 'identity')

#表格

table(Churn)

#用户属性：gender,SeniorCitizen,Partner,Dependents,tenure

Percentage<- matrix(rep(1,nrow(data)),nrow=(nrow(data)),ncol=1)

plot1<-ggplot(data, aes(x=gender, y=Percentage, fill=Churn))+geom\_col(position="fill")

plot2<-ggplot(data, aes(x=SeniorCitizen, y=Percentage, fill=Churn))+geom\_col(position="fill")

plot3<-ggplot(data, aes(x=Partner, y=Percentage, fill=Churn))+geom\_col(position="fill")

plot4<-ggplot(data, aes(x=Dependents, y=Percentage, fill=Churn))+geom\_col(position="fill")

grid.newpage()

pushViewport(viewport(layout = grid.layout(2,2)))

vplayout <- function(x,y)

viewport(layout.pos.row = x, layout.pos.col = y)

print(plot1, vp = vplayout(1, 1))

print(plot2, vp = vplayout(1, 2))

print(plot3, vp = vplayout(2, 1))

print(plot4, vp = vplayout(2, 2))

#服务属性：PhoneService,MultipleLines,InternetService,OnlineSecurity,OnlineBackup,

#DeviceProtection,TechSupport,StreamingTV,StreamingMovies

#流失率对比

plot4<-ggplot(data, aes(x=PhoneService, y=Percentage, fill=Churn))+geom\_col(position="fill")

plot5<-ggplot(data, aes(x=MultipleLines, y=Percentage, fill=Churn))+geom\_col(position="fill")

plot6<-ggplot(data, aes(x=InternetService, y=Percentage, fill=Churn))+geom\_col(position="fill")

plot7<-ggplot(data, aes(x=OnlineSecurity, y=Percentage, fill=Churn))+geom\_col(position="fill")

plot8<-ggplot(data, aes(x=OnlineBackup, y=Percentage, fill=Churn))+geom\_col(position="fill")

plot9<-ggplot(data, aes(x=DeviceProtection, y=Percentage, fill=Churn))+geom\_col(position="fill")

plot10<-ggplot(data, aes(x=TechSupport, y=Percentage, fill=Churn))+geom\_col(position="fill")

plot11<-ggplot(data, aes(x=StreamingTV, y=Percentage, fill=Churn))+geom\_col(position="fill")

plot12<-ggplot(data, aes(x=StreamingMovies, y=Percentage, fill=Churn))+geom\_col(position="fill")

grid.newpage()

# pushViewport函数提供了添加视图以及在树中的视图之间导航的方法。

pushViewport(viewport(layout = grid.layout(3,3)))

# viewport函数创建视图，描述图形设备上的矩形区域，并在这些区域中定义许多坐标系统。

vplayout <- function(x,y)

viewport(layout.pos.row = x, layout.pos.col = y)

print(plot4, vp = vplayout(1, 1))

print(plot5, vp = vplayout(1, 2))

print(plot6, vp = vplayout(1, 3))

print(plot7, vp = vplayout(2, 1))

print(plot8, vp = vplayout(2, 2))

print(plot9, vp = vplayout(2, 3))

print(plot10, vp = vplayout(3, 1))

print(plot11, vp = vplayout(3, 2))

print(plot12, vp = vplayout(3, 3))

#合同属性：MonthlyCharges,TotalCharges,Contract,PaperlessBilling,PaymentMethod

#相关性分析：1.分类变量2.连续变量

#1.分类变量

myFUN<- function(x){chisq.test(data$Churn,x ,correct = TRUE)}

#apply（数据库，循环数据库3到6列，按列，函数）

colnames(data)

xx = c("customerID","tenure","MonthlyCharges","TotalCharges")

dataset = data[,!names(data) %in% xx]

result<- apply (dataset,2, myFUN)

#创建功能，提取list每个循环的p.value

p<- function(x){x$p.value}

#转化成数据，sapple(数据框，函数)

#添加比较组的行名

A<-names(dataset)

A

#去掉不相关变量

xx = c("customerID","gender","TotalCharges")

data1 = data[,!names(data) %in% xx]

head(data1)

科技发展与生态保护

2113850 李鹏 软件工程

近几十年来，我国科技发展迅速，人们生活水平不断提高，但与此同时，也给环境带来了很大的破坏。生态环境现状不容乐观，能源资源面临枯竭，如何在加快发展科学技术的同时，保护好生态环境，成为我们迫在眉睫的研究课题。

当然，随着自动化、物联网、大数据等技术的快速发展，越来越多的先进的算法、软件工具已经可以处理海量的数据信息，给生态文明保护开辟了新方向，带来了新希望。随着人工智能技术应用领域的日益扩大，人工智能技术通过优化机器性能、减少能源消耗、控制污染物生成、提高检测效率等方式，不断地为环境保护工作注入新的活力。软件开发在其中发挥着不言而喻的作用。

生态环境的保护迫在眉睫，信息化的发展也为生态保护提供了“数字环保”的新理念。现阶段，随着我国信息化建设的进一步深入发展与完善，以信息化为特色的环保措施陆续出台，取得成效。人工智能技术的出现和应用使得环保领域又增添了新的活力，成为环保工作的强大技术保证。

我相信，在不久的将来，以信息化为主导的生态环境保护将成为生态文明保护的主力军。另外，最重要的是，我们每个人也都要养成生态环境保护的理念，并付诸实践，从根源上解决生态环境保护问题。