# 加载必要的包

library(readxl) # 用于读取Excel文件

library(caret) # 用于模型训练和评估

library(randomForest) # 随机森林模型

library(writexl) # 用于导出Excel文件

# 读取数据

trainData <- read\_excel("C:/Users/Liu/Desktop/13ta.xlsx")

testData <- read\_excel("C:/Users/Liu/Desktop/13te.xlsx")

# 设定种子以确保可重复性

set.seed(3)

# 检查数据结构

str(trainData)

str(testData)

# 检查训练数据中的缺失值

missing\_train <- sum(is.na(trainData))

cat("训练数据集中的缺失值总数：", missing\_train, "\n")

# 检查测试数据中的缺失值

missing\_test <- sum(is.na(testData))

cat("测试数据集中的缺失值总数：", missing\_test, "\n")

# 处理缺失值（这里采用均值填补，但也可以用其他方法）

# 确保测试数据中的特征列与训练数据一致

testData <- testData[ , colnames(trainData)]

# 设置训练控制

controlD <- trainControl(method = "cv", number = 10, search = "grid")

# 训练随机森林模型

rf\_model <- train(

AGB ~ .,

data = trainData,

method = 'rf',

trControl = controlD

)

# 打印模型以查看结果

print(rf\_model)

# 进行预测

predictions <- predict(rf\_model, newdata = testData)

# 将预测结果与实际标签（如果有）导出到 Excel 文件

# 如果测试数据集中有真实标签，请确保它们的列名为 'AGB'

if ("AGB" %in% colnames(testData)) {

results <- data.frame(Actual = testData$AGB, Predicted = predictions)

} else {

results <- data.frame(Predicted = predictions)

}

# 导出结果到 Excel 文件

write\_xlsx(results, "C:/Users/asus/Desktop/15RF.xlsx")

cat("预测结果已导出到 'C:/Users/asus/Desktop/predictions\_results.xlsx'\n")