SIMPLE REGRESSION MODEL TO PREDICT MONTHLY POCKET MONEY EXPENSES

A PROJECT REPORT

Submitted by

NEKASHRI S(2303811724322076)

in partial fulfillment of requirements for the award of the course

AGI1252 - FUNDAMENTALS OF DATA SCIENCE USING R

in

ARTIFICIAL INTELLIGENCE AND DATA SCIENCE

K. RAMAKRISHNAN COLLEGE OF TECHNOLOGY

(An Autonomous Institution, affiliated to Anna University Chennai and Approved by AICTE, New Delhi)

SAMAYAPURAM – 621 112

JUNE- 2025

K. RAMAKRISHNAN COLLEGE OF TECHNOLOGY (AUTONOMOUS)

SAMAYAPURAM – 621 112

BONAFIDE CERTIFICATE

Certified that this project report on **" SIMPLE REGRESSION MODEL TO PREDICT MONTHLY POCKET MONEY EXPENSES "** is the bonafide work of **NEKASHRI S(230381172432207 6)**who carried out the project work during the academic year 2024 - 2025 under my supervision.

SIGNATURE

Dr.T. AVUDAIAPPAN, M.E.,Ph.D.,

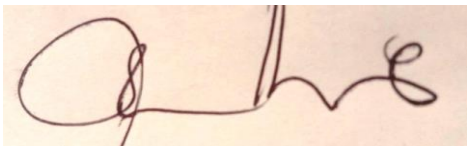
HEAD OF THE DEPARTMENT

PROFESSOR

Department of Artificial Intelligence

K.Ramakrishnan College of Technology (Autonomous)

Samayapuram–621112.



SIGNATURE

Ms.S.Murugavalli., M.E.,(Ph.D).,

SUPERVISOR

ASSISTANT PROFESSOR

Department of Artificial Intelligence

K.Ramakrishnan College of Technology (Autonomous)

Samayapuram–621112.

Submitted for the viva-voce examination held on **02.06.2025**.

INTERNAL EXAMINER

EXTERNAL EXAMINER

DECLARATION

I declare that the project report on **"SIMPLE REGRESSION MODEL TO PREDICT MONTHLY POCKET MONEY EXPENSES "** is the result of original work done by us and best of our knowledge, similar work has not been submitted to **"ANNA UNIVERSITY CHENNAI"** for the requirement of Degree of **BACHELOR OF TECHNOLOGY**. This project report is submitted on the partial fulfilment of the requirement of the completion of the course **AGI1252 - FUNDAMENTALS OF DATA SCIENCE USING R**

.

Signature

NEKASHRI S

Place: Samayapuram

Date:2.06.2025

ACKNOWLEDGEMENT

It is with great pride that I express our gratitude and in-debt to our institution "**K.Ramakrishnan College of Technology (Autonomous)**", for providing us with the opportunity to do this project.

I glad to credit honourable chairman **Dr. K. RAMAKRISHNAN**, **B.E.,** for having provided for the facilities during the course of our study in college.

I would like to express our sincere thanks to our beloved Executive Director **Dr. S. KUPPUSAMY, MBA, Ph.D.,** for forwarding to our project and offering adequate duration in completing our project.

I would like to thank **Dr. N. VASUDEVAN, M.Tech., Ph.D.,** Principal, who gave opportunity to frame the project the full satisfaction.

I whole heartily thanks to **Dr. T. AVUDAIAPPAN, M.E.,Ph.D.,** Head of the department, **ARTIFICIAL INTELLIGENCE** for providing his encourage pursuing this project.

I express our deep expression and sincere gratitude to our project supervisor **Ms.S.Murugavalli., M.E.,(Ph.D).,** Department of **ARTIFICIAL INTELLIGENCE,** for her incalculable suggestions, creativity, assistance and patience which motivated us to carry out this project.

I render our sincere thanks to Course Coordinator and other staff members for providing valuable information during the course.

I wish to express our special thanks to the officials and Lab Technicians of our departments who rendered their help during the period of the work progress.

INSTITUTE

Vision:

- To serve the society by offering top-notch technical education on par with global standards.

Mission:

- Be a center of excellence for technical education in emerging technologies by exceeding the needs of industry and society.
- Be an institute with world class research facilities.
- Be an institute nurturing talent and enhancing competency of students to transform them as all – round personalities respecting moral and ethical values.

DEPARTMENT

Vision:

- To excel in education, innovation, and research in Artificial Intelligence and Data Science to fulfil industrial demands and societal expectations.

Mission

- To educate future engineers with solid fundamentals, continually improving teaching methods using modern tools.
- To collaborate with industry and offer top-notch facilities in a conducive learning environment.
- To foster skilled engineers and ethical innovation in AI and Data Science for global recognition and impactful research.
- To tackle the societal challenge of producing capable professionals by instilling employability skills and human values.

PROGRAM EDUCATIONAL OBJECTIVES (PEO)

- **PEO1:** Compete on a global scale for a professional career in Artificial Intelligence and Data Science.
- **PEO2:** Provide industry-specific solutions for the society with effective communication and ethics.
- **PEO3** Enhance their professional skills through research and lifelong learning initiatives.

PROGRAM SPECIFIC OUTCOMES (PSOs)

- **PSO1:** Capable of finding the important factors in large datasets, simplify the data, and improve predictive model accuracy.
- **PSO2:** Capable of analyzing and providing a solution to a given real-world problem by designing an effective program.

PROGRAM OUTCOMES (POs)

Engineering students will be able to:

1. **Engineering knowledge:** Apply knowledge of mathematics, natural science, computing, engineering fundamentals, and an engineering specialization to develop solutions to complex engineering problems.

2. **Problem analysis:** Identify, formulate, review research literature and analyze complex engineering problems reaching substantiated conclusions with consideration for sustainable development.

3. **Design/development of solutions:** Design creative solutions for complex engineering problems and design/develop systems/components/processes to meet identified needs with consideration for the public health and safety, whole-life cost, net zero carbon, culture, society and environment as required.

4. **Conduct investigations of complex problems:** Conduct investigations of complex engineering problems using research-based knowledge including design of experiments, modelling, analysis & interpretation of data to provide valid conclusions.

5. **Engineering Tool Usage:** Create, select and apply appropriate techniques, resources and modern engineering & IT tools, including prediction and modelling recognizing their limitations to solve complex engineering problems.

6. **The Engineer and The World:** Analyze and evaluate societal and environmental aspects while solving complex engineering problems for its impact on sustainability with reference to economy, health, safety, legal framework, culture and environment.

7. **Ethics:** Apply ethical principles and commit to professional ethics, human values, diversity and inclusion; adhere to national & international laws.

8. **Individual and Collaborative Team work:** Function effectively as an individual, and as a member or leader in diverse/multi-disciplinary teams.

9. **Communication:** Communicate effectively and inclusively within the engineering community and society at large, such as being able to comprehend and write effective reports and design documentation, make effective presentations considering cultural, language, and learning differences.

10. **Project management and finance:** Apply knowledge and understanding of engineering management principles and economic decision-making and apply these to one's own work, as a member and leader in a team, and to manage projects and in multidisciplinary environments.

11. **Life-long learning:** Recognize the need for, and have the preparation and ability for i) independent and life-long learning ii) adaptability to new and emerging technologies and iii) critical thinking in the broadest context of technological change.

ABSTRACT

A simple linear regression model was developed using R to predict monthly pocket money expenses based on influencing variables such as spending habits or itemized costs. The implementation leverages R's Shiny framework to create an interactive web-based application that allows users to upload CSV datasets, select numeric variables, and visualize relationships through scatter plots and summary statistics. Object-Oriented Programming principles, implemented via R6 classes, streamline data handling and ensure modular code structure. The model facilitates understanding of how different factors affect spending, making it a valuable tool for students, researchers, and financial planners aiming to forecast or manage personal expenses effectively. Interactive customization features such as point size and color selection enhance user experience and visualization clarity, while extensive testing with varied datasets ensures robustness and flexibility in real-world usage scenarios.

ABSTRACT WITH POs AND PSOs MAPPING

CO 5 : BUILD DATA SCIENCE USING R PROGRAMMING FOR SOLVING REAL-TIME PROBLEMS.

ABSTRACT	POs MAPPED	PSOs MAPPED
A linear regression model using R and Shiny predicts monthly pocket money expenses based on Income, Age, and Savings. The app allows CSV uploads, visualizes data with scatter plots, and displays model summaries and evaluation metrics. Object-Oriented Programming (R6) improves data handling. With customization and testing features, it serves as a reliable tool for financial analysis and learning.	**PO1 -3** **PO2 -3** **PO3 -3** **PO4 -3** **PO5 -3** **PO6 -3** **PO7 -3** **PO8 -3** **PO9 -3** **PO10 -3** **PO11-3**	**PSO1 -3** **PSO2 -3**

Note: 1- Low, 2-Medium, 3- High

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

1.1 Objective

The primary objective of this project is to develop an interactive web application that predicts monthly pocket money expenses based on key personal financial factors such as income, age, and savings. The application allows users to upload financial datasets and builds a predictive model that reveals the relationships between these variables and spending behavior. By providing clear visualizations and easy-to-understand statistical summaries, the tool aims to help students, researchers, and financial planners gain deeper insights into personal expense patterns. Ultimately, the application serves as a practical aid for forecasting expenses and supporting better financial management and decision-making.

1.2 Overview

This project presents an interactive platform for analyzing and predicting monthly pocket money expenses using statistical modeling and data visualization. Users can upload datasets containing relevant financial variables, which the platform processes to ensure data integrity before constructing a predictive model. The core modeling technique establishes a relationship between income, age, savings, and expenses, allowing users to explore how these factors influence spending.

The application offers multiple functionalities, including scatter plots that visually depict correlations between variables, comprehensive model summaries detailing coefficient estimates and significance levels, and evaluation metrics that quantify model accuracy and fit. Users are also able to input custom data values to generate personalized expense predictions, enhancing the tool's practical usefulness.

Designed with a focus on user-friendliness and interactivity, the platform enables meaningful exploration of financial data even for those without advanced statistical expertise.

1.3 Data Science related concepts

- **Interactive Web Development:** Utilizes a reactive programming framework that enables the creation of dynamic user interfaces. This allows the application to respond in real time to data uploads, user inputs, and button clicks without page reloads.
- **Statistical Modeling:** Implements linear regression to establish and quantify the relationship between monthly expenses and predictor variables such as income, age, and savings. This technique models how changes in these factors affect spending behavior.
- **Data Validation:** Ensures that uploaded datasets contain all necessary variables before proceeding with analysis, preventing errors and enhancing reliability.
- **Data Visualization:** Employs advanced plotting libraries to generate clear and informative scatter plots with regression trend lines, aiding users in visually interpreting data relationships.
- **User Input Controls and Event Handling:** Provides interactive numeric input fields and action buttons that allow users to supply new data points and trigger expense predictions, fostering engagement and customization.
- **Model Performance Assessment:** Calculates evaluation metrics such as Root Mean Square Error (RMSE) and R-squared values, giving users quantitative feedback on model accuracy and goodness-of-fit.
- **Reactive Expressions and Observers:** Manage the flow of data and computations efficiently by automatically updating outputs whenever inputs or datasets change, ensuring seamless user experience.

CHAPTER 2

PROJECT METHODOLOGY

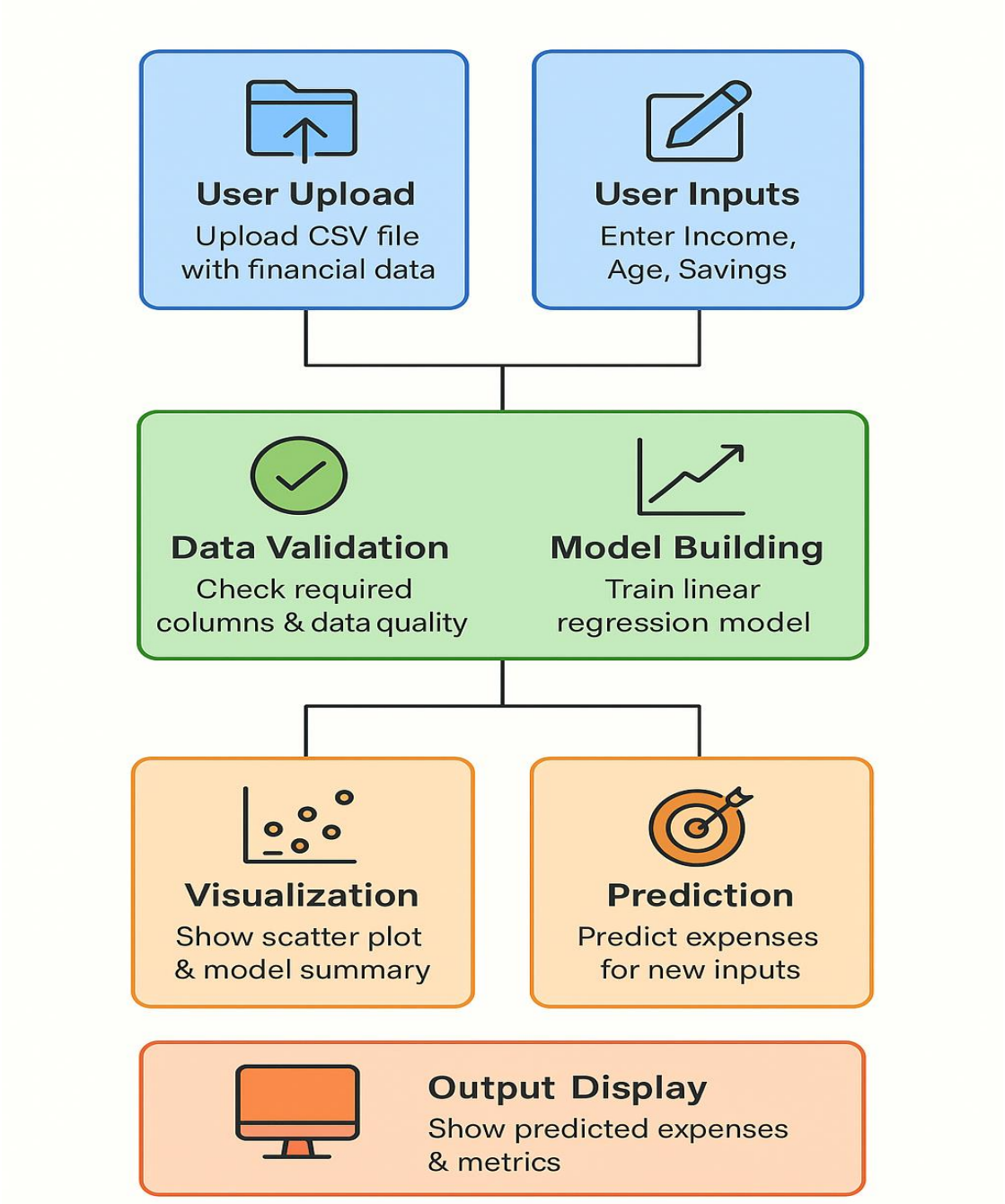
2.1 Proposed Work

The proposed work involves developing an interactive application that enables users to predict monthly pocket money expenses based on their personal financial information. The application accepts a dataset containing historical financial records, including income, age, savings, and expenses, and performs data validation to ensure its integrity. Using this data, a linear regression model is constructed to identify and quantify the relationship between these factors and monthly expenses.

The application features a user-friendly interface that allows users to visualize data trends through scatter plots, examine detailed statistical summaries of the regression model, and evaluate model performance using key metrics such as RMSE and R-squared. Furthermore, users can input new financial data values to obtain personalized expense predictions in real time.

This methodology focuses on delivering an easy-to-use and dynamic tool that supports data-driven financial decision-making and personal expense management. The interactive nature of the application encourages experimentation with different input variables, providing insights into how various factors influence spending behavior.

2.2 Block Diagram



CHAPTER 3

MODULE DESCRIPTION

3.1 Data Upload Module

The Data Upload module serves as the starting point where users upload their CSV datasets containing financial information such as Income, Age, Savings, and Expenses. The interface provides a file input control that restricts uploads to CSV files only, ensuring data consistency. Upon upload, the system validates the dataset to confirm all necessary columns are present, preventing errors during analysis. This module ensures users work with clean, reliable data, which is essential for accurate model training and predictions.

3.2 Input and Prediction Module

This module allows users to enter new financial details—Income, Age, and Savings—using numeric input fields. Once users provide these values and click the “Predict Expenses” button, the module processes the inputs through the trained linear regression model. It outputs a predicted monthly expense value based on the user’s inputs. This interactive prediction functionality enables personalized expense forecasting, helping users understand how changes in their financial variables might affect spending.

3.3 Data Visualization Module

The Data Visualization module presents insightful graphical representations of the uploaded data. It includes a scatter plot that displays the relationship between Income and Expenses, complemented by a regression trend line to highlight the model fit. These visual tools help users easily grasp correlations within their data. Interactive and responsive, the visualization updates automatically whenever new data is uploaded, allowing dynamic exploration of spending behavior trends.

3.4 Model Summary and Evaluation Module

This module provides detailed statistical information about the linear regression model built from the uploaded data. It displays a comprehensive summary including coefficients, standard errors, and significance levels, giving users an understanding of the model parameters. Additionally, key performance metrics such as Root Mean Square Error (RMSE) and R-squared values are shown to assess model accuracy and goodness-of-fit. This information helps users evaluate how well the model explains the variance in monthly expenses.

3.5 User Interface and Interaction Module

The User Interface module ensures a seamless and intuitive experience for users navigating the application. It integrates all input controls, buttons, and output displays into a cohesive layout. Event handlers listen for user actions such as file uploads or prediction requests and update the interface reactively without page reloads. This module prioritizes usability and accessibility, making the application suitable for users with varying levels of statistical and technical expertise.

CHAPTER 4

CONCLUSION & FUTURE SCOPE

4.1 Conclusion

The Pocket Money Expense Predictor effectively utilizes linear regression to forecast monthly expenses based on key financial variables such as income, age, and savings. Developed using R and the Shiny framework, the application offers an interactive and user-friendly platform where users can upload their data, visualize spending trends, and receive personalized expense predictions. The combination of data validation, model evaluation metrics, and dynamic visualizations ensures accuracy and transparency, making it a valuable tool for individuals aiming to better manage their personal finances.

Its modular and reactive design encourages users to understand how different financial factors influence their spending, supporting informed decision-making and enhanced financial planning. Overall, the system provides a practical approach to monitoring and predicting pocket money expenses, particularly beneficial for students and young adults.

4.2 Future Scope

- Incorporating additional variables such as lifestyle habits and geographic factors to improve prediction quality.
- Employing advanced machine learning techniques like Random Forest or Gradient Boosting to capture complex spending patterns.
- Adding secure user authentication and encrypted data storage to protect sensitive information.
- Developing a mobile-friendly version or standalone app for on-the-go expense tracking.
- Introducing budgeting tools, spending alerts, and saving recommendations to aid financial discipline.
- Offering multi-language support and customizable UI themes to enhance accessibility and user experience.

CHAPTER 5

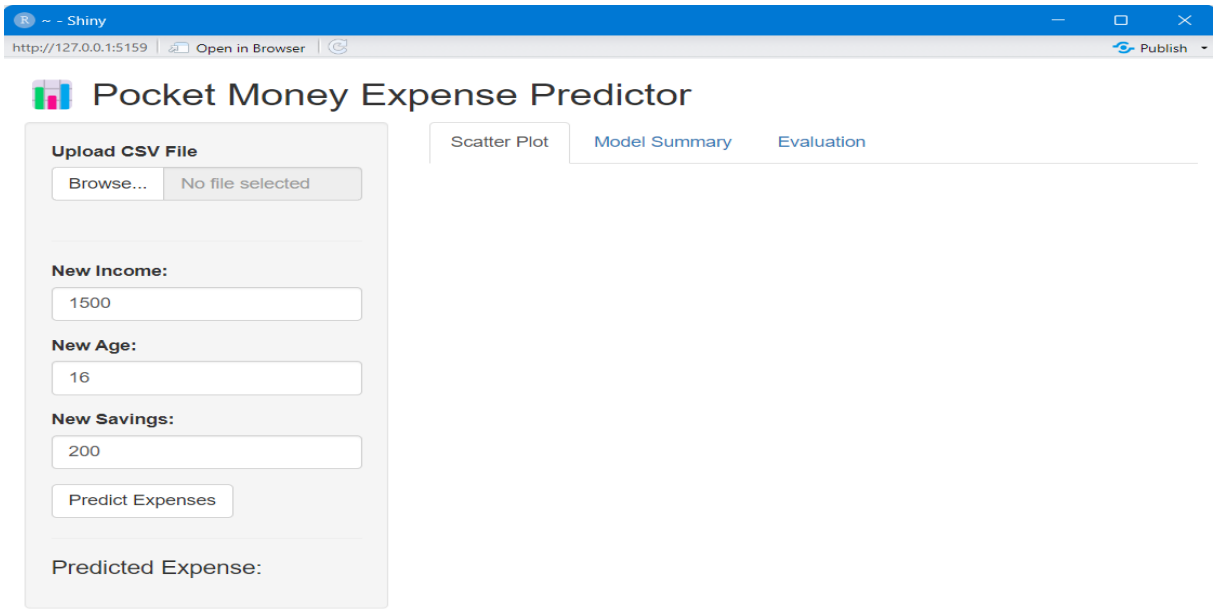
APPENDIX A – SOURCE CODE

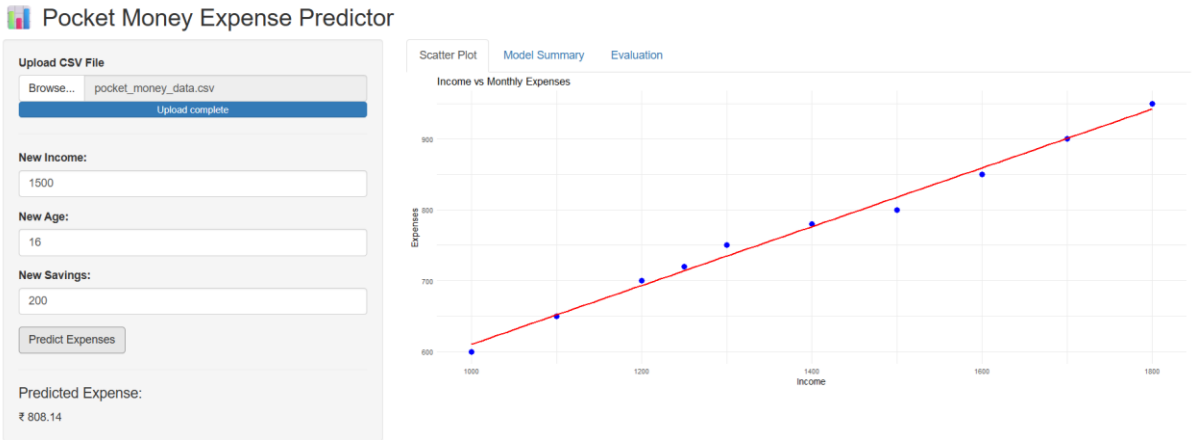
```
library(shiny)
library(ggplot2)
library(dplyr)

ui <- fluidPage(
  titlePanel("📊 Pocket Money Expense Predictor"),

  sidebarLayout(
    sidebarPanel(
      fileInput("file", "Upload CSV File", accept = ".csv"),
      tags$hr(),
      numericInput("income", "New Income:", value = 1500),
      numericInput("age", "New Age:", value = 16),
      numericInput("savings", "New Savings:", value = 200),
      actionButton("predict", "Predict Expenses"),
      tags$hr(),
      h4("Predicted Expense:"),
      textOutput("prediction")
    ),

    mainPanel(
      tabsetPanel(
        tabPanel("Scatter Plot", plotOutput("scatterPlot")),
        tabPanel("Model Summary", verbatimTextOutput("modelSummary")),
        tabPanel("Evaluation", verbatimTextOutput("metrics"))
      )
```

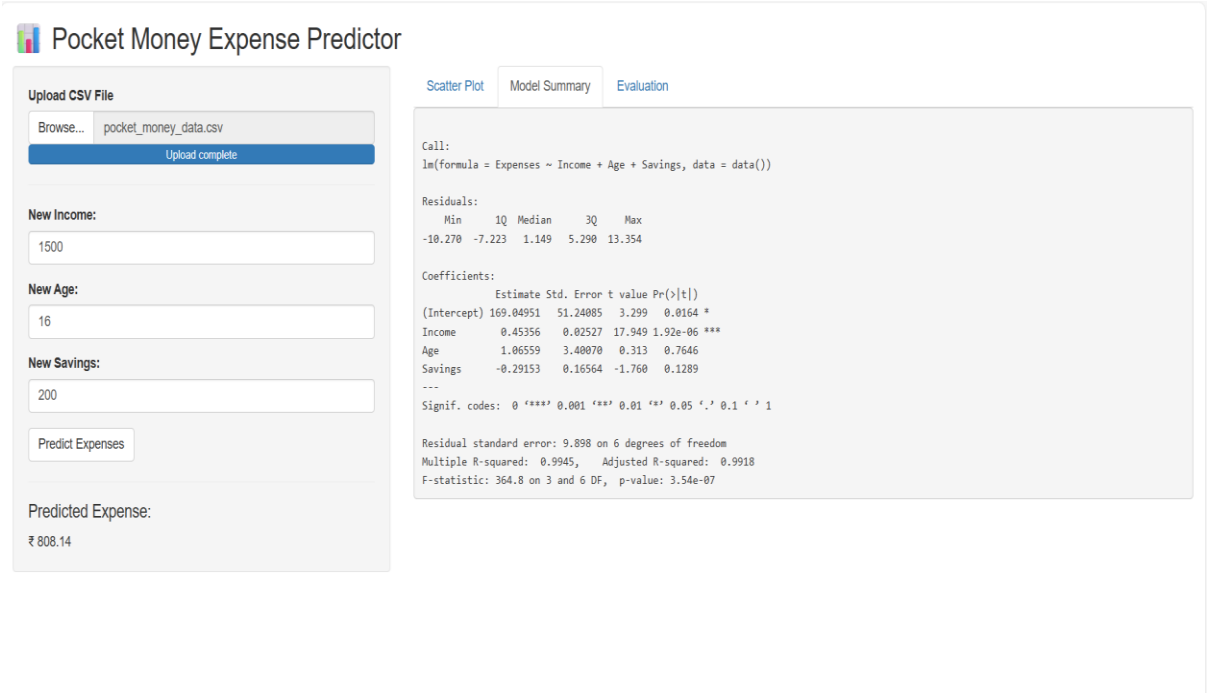
```
    )
  )
)

server <- function(input, output) {

  data <- reactive({
    req(input$file)
    df <- read.csv(input$file$datapath)
    validate(
      need(all(c("Income", "Age", "Savings", "Expenses") %in% names(df)),
           "CSV must contain: Income, Age, Savings, Expenses")
    )
    return(df)
  })

  model <- reactive({
    req(data())
    lm(Expenses ~ Income + Age + Savings, data = data())
  })

  output$scatterPlot <- renderPlot({
    req(data())
    ggplot(data(), aes(x = Income, y = Expenses)) +
      geom_point(color = "blue", size = 3) +
      geom_smooth(method = "lm", se = FALSE, color = "red") +
      ggtitle("Income vs Monthly Expenses") +
      theme_minimal()
  })
```

```
  output$modelSummary <- renderPrint({
    req(model())
    summary(model())
  })

  output$metrics <- renderPrint({
    req(model())
    preds <- predict(model(), data())
    rmse <- sqrt(mean((preds - data()$Expenses)^2))
    r2 <- summary(model())$r.squared
    cat("RMSE:", round(rmse, 2), "\n")
    cat("R-squared:", round(r2, 2), "\n")
  })
  observeEvent(input$predict, {
    req(model())
    new_user <- data.frame(
      Income = input$income,
      Age = input$age,
      Savings = input$savings
    )
    pred <- predict(model(), new_user)
    output$prediction <- renderText({
      paste("₹", round(pred, 2))
    })
  })
}

shinyApp(ui = ui, server = server)
```

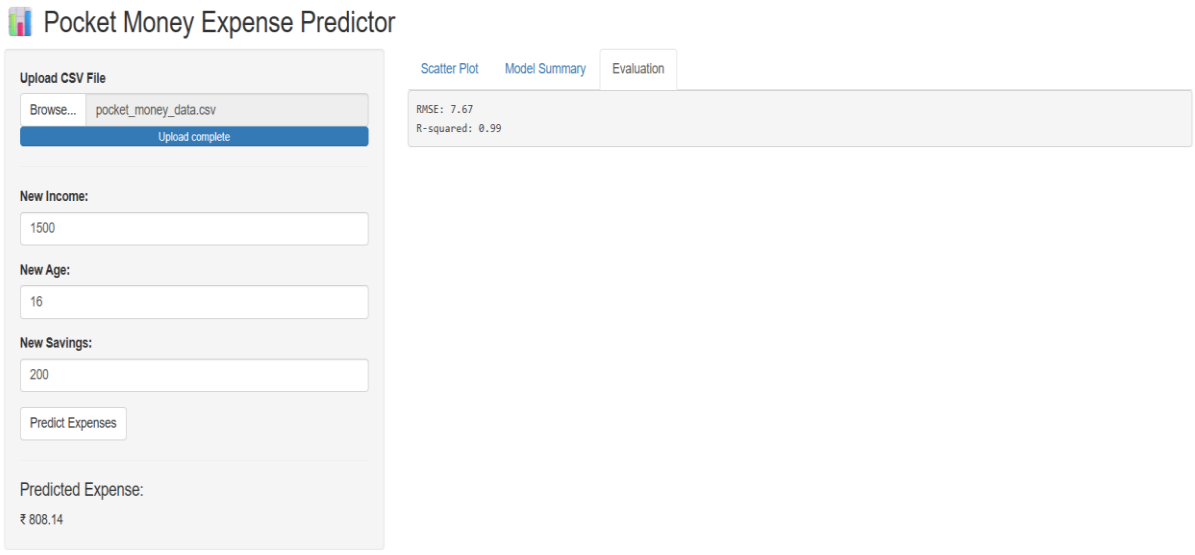
Appendix B – Screenshots






Pocket Money Expense Predictor
Upload CSV File
Browse...
pocket_money_data.csv
Upload complete
New Income:
1500
New Age:
16
New Savings:
200
Predict Expenses
Predicted Expense:
₹ 808.14
Scatter Plot
Model Summary
Evaluation
Call:
lm(formula = Expenses ~ Income + Age + Savings, data = data())
Residuals:
Min 1Q Median 3Q Max
-10.270 -7.223 1.149 5.290 13.354
Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept) 169.04951 51.24085 3.299 0.0164 *
Income 0.45356 0.02527 17.949 1.92e-06 ***
Age 1.06559 3.40070 0.313 0.7646
Savings -0.29153 0.16564 -1.760 0.1289

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
Residual standard error: 9.898 on 6 degrees of freedom
Multiple R-squared: 0.9945, Adjusted R-squared: 0.9918
F-statistic: 364.8 on 3 and 6 DF, p-value: 3.54e-07


Pocket Money Expense Predictor
Upload CSV File
Browse...
pocket_money_data.csv
Upload complete
New Income:
1500
New Age:
16
New Savings:
200
Predict Expenses
Predicted Expense:
₹ 808.14
Scatter Plot
Model Summary
Evaluation
RMSE: 7.67
R-squared: 0.99

REFERENCES

1. R Core Team (2024). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. Available at: https://www.r-project.org/
2. Wickham, H. et al. (2023). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag. Available at: https://ggplot2.tidyverse.org/
3. Chang, W. et al. (2024). *shiny: Web Application Framework for R*. R package version 1.8.0. Available at: https://shiny.rstudio.com/
4. Wickham, H. et al. (2023). *dplyr: A Grammar of Data Manipulation*. R package version 1.1.4. Available at: https://dplyr.tidyverse.org/
5. Hadley Wickham and Lionel Henry (2024). *tidyr: Tidy Messy Data*. R package version 1.3.0. Available at: https://tidyr.tidyverse.org/
6. Kabacoff, R. (2021). *R in Action: Data Analysis and Graphics with R*, Second Edition. Manning Publications.