
# DAILY WORK REPORT

# TR-02

## INFOWIZ

20 JUNE 2024

# Day 14: Project - Implementing Linear Regression

**Summary:** Today marked the beginning of our project phase where we applied our learning from previous weeks to implement linear regression on a real-world dataset. This hands-on project aimed to reinforce our understanding of linear regression concepts, data preprocessing techniques, and model evaluation.

**Key Activities:**

1. **Project Selection:**
    - Brainstormed and selected a suitable dataset for our project, considering factors like data size, complexity, and relevance to linear regression analysis.
    - Chose a dataset related to housing prices to predict home values based on features such as square footage, number of bedrooms, and location.
2. **Data Exploration and Preprocessing:**
    - Loaded and explored the dataset using Pandas, examining key statistics, data types, and initial insights into feature distributions.
    - Conducted data cleaning tasks such as handling missing values, encoding categorical variables, and scaling numerical features using techniques learned in previous sessions with NumPy and Pandas.
3. **Implementing Linear Regression:**
    - Applied scikit-learn's LinearRegression class to fit the linear regression model to our preprocessed dataset.
    - Evaluated model performance using metrics like mean squared error (MSE) and R-squared to assess the model's ability to predict housing prices accurately.
4. **Documentation and Reporting:**
    - Documented the project methodology, including data preprocessing steps, model implementation details, and evaluation metrics.
    - Summarized findings and insights derived from the linear regression analysis, highlighting key predictors influencing housing prices and any challenges encountered during the project.

**Next Steps:**

- Refine the linear regression model by experimenting with different features and model parameters to improve predictive accuracy.
- Explore additional regression techniques such as Ridge Regression or Lasso Regression to address potential overfitting and enhance model robustness.
- Prepare for the final presentation by organizing project documentation, creating visual aids, and rehearsing the presentation to effectively communicate project outcomes and insights.

Today's project implementation phase provided practical experience in applying linear regression to real-world data, reinforcing theoretical concepts learned throughout the course and setting the stage for deeper exploration into advanced machine learning techniques in the coming weeks.