DSA210 Final Report

The Impact of Smartphone Notifications on Cognitive Control

Student: Metin Gündüz
Course: DSA 210 – Introduction to Data Science
Instructor: Özgür Asar
Term: Spring 2025

---

1. Introduction

Smartphones generate hundreds of notifications per day, fragmenting attention and potentially degrading executive control. This project investigates how daily notification load correlates with performance on a Go/No-Go task, a standard test of response inhibition. We additionally assess the roles of mood and caffeine intake.

---

2. Data Collection & Description

- **Period**: 12 March – 25 April 2025 (45 days)
- **Device**: iPhone (notification tracking via Screen Time)
- **Task**: Browser-based Go/No-Go task (PsyToolkit)
- **Variables**:
    - Notifications (daily count)
    - Commission Error Rate (%)
    - Reaction Time (ms)
    - Mood (1–5 scale)
    - Caffeine intake (mg)

Data were collected manually and semi-automatically, converted from JSON to CSV using Google Colab, and merged on date. Cleaning steps included numeric conversions, character stripping, and null removal.

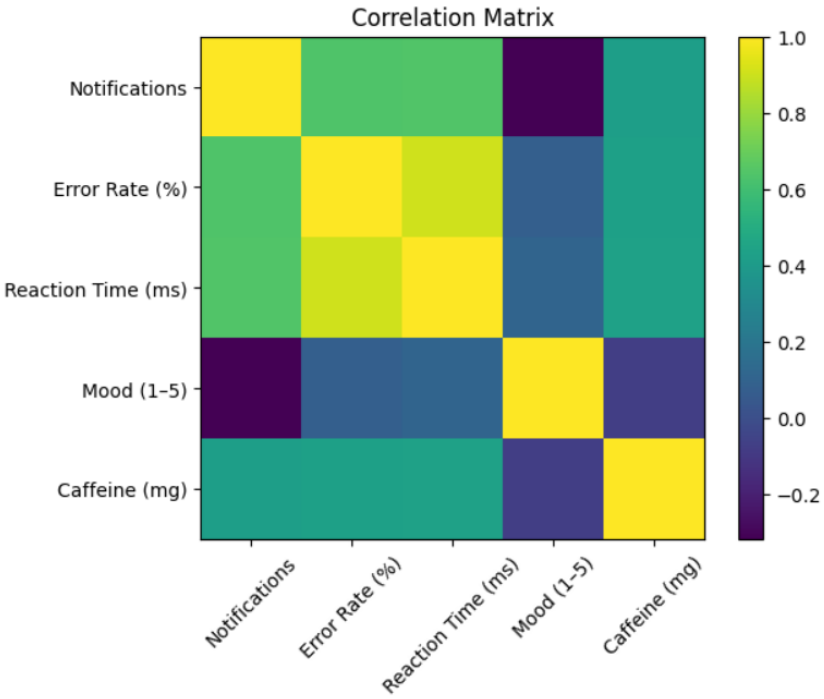
---

3. Exploratory Data Analysis (EDA)

3.1 Distributions

- Notifications mostly ranged between 250–400.
- Error Rate between 8–18%.
- Reaction Time 450–600ms.
- Mood typically 3–4 with notable dips.
- Caffeine peaked at 0, 60, 120, and 180mg.

3.2 Correlation Matrix

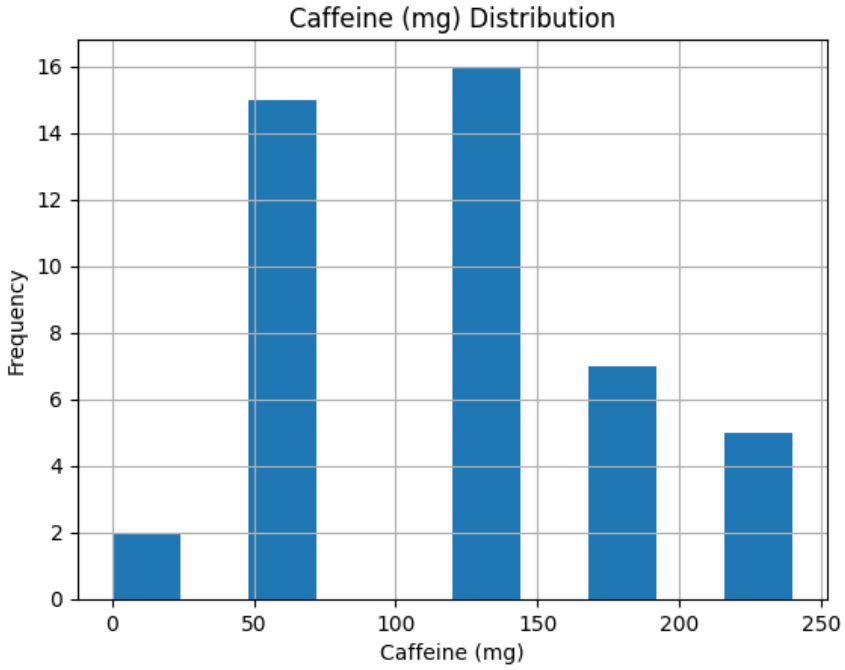


- Strong correlation between Error Rate and Notifications (r ≈ 0.64).
- Error Rate and Reaction Time also strongly correlated (r ≈ 0.90).

3.3 Key Scatter Plots

Notifications vs. Error Rate

Caffeine vs. Reaction Time



---

4. Hypothesis Testing

- **H0**: There is no relationship between notification count and error rate.
- **H1**: There is a positive linear relationship.

Result: Pearson r = 0.637, p < 0.001 → Reject H0.
Conclusion: High notification days are associated with more commission errors.

Additional:

- **Mood vs. Error Rate**: r = 0.08 (p = 0.58) → No correlation.
- **Caffeine vs. Reaction Time**: r = 0.436 (p = 0.003) → Moderate correlation.

---

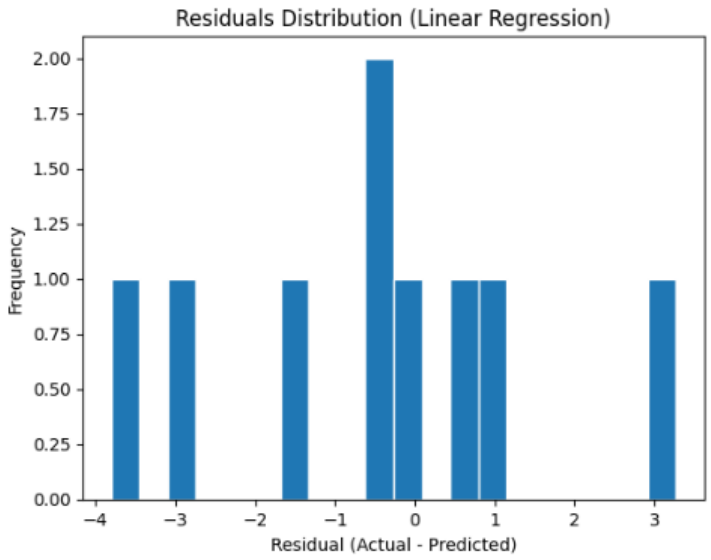
5. Machine Learning Modeling

5.1 Regression (Predicting Error Rate)

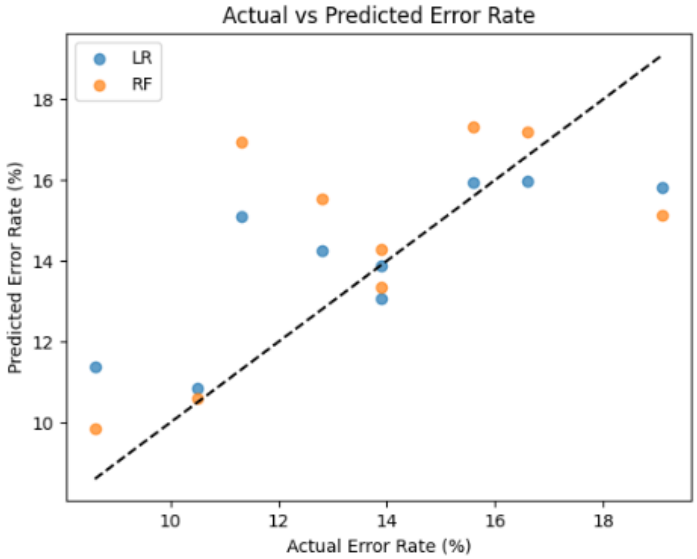
Features: Notifications, Mood, Caffeine
Models:

Model	MSE	R^2
Linear Regression	4.04	0.57
Random Forest	6.72	0.28

Residual Histogram (LR)



Actual vs. Predicted (LR vs. RF)

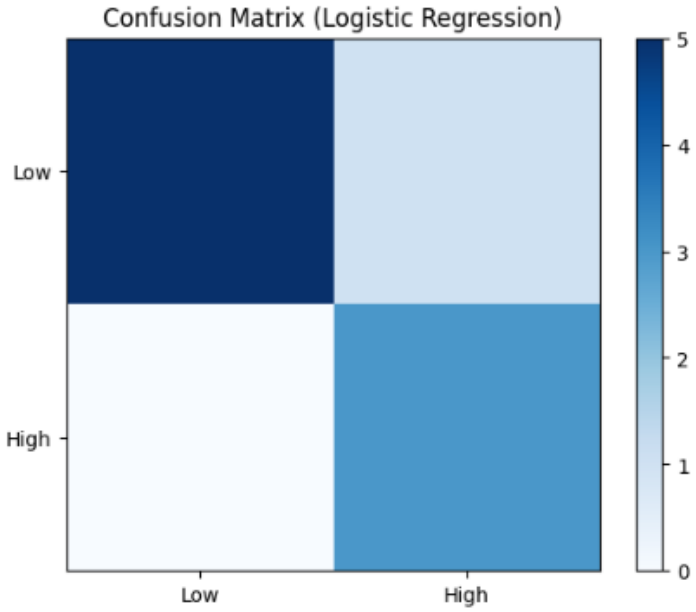


5.2 Classification (High vs Low Error)

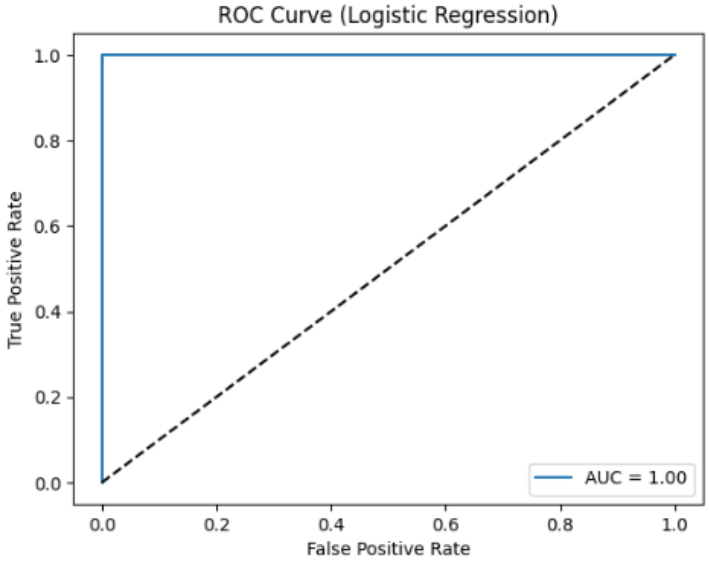
Labeling: Error Rate > median = High Error Day
Models: Logistic Regression, Decision Tree

Model	Accuracy	AUC
Logistic Regression	0.89	1.00
Decision Tree	0.78	0.88

Confusion Matrix (LogReg)



ROC Curve (LogReg)



---

6. Interpretation & Discussion

- **Notifications** was consistently the most predictive feature across tasks.
- **Caffeine** had a moderate but unexpected positive correlation with slower reaction times.
- **Mood** had weak or inconsistent predictive value.
- ML models performed well despite the small dataset; Logistic Regression classified high-error days with 0.89 accuracy and perfect AUC.

---

7. Limitations & Future Work

- Sample size limited to 45 days.
- Mood and caffeine self-reported; sleep and stress were not tracked.
- Future work could:
    1. Expand to 90+ days
    2. Incorporate sleep metrics and screen time types
    3. Use cross-validation and hyperparameter tuning
    4. Apply time-lagged models

---

8. Conclusion

This project supports the hypothesis that frequent daily smartphone notifications negatively affect cognitive control. Notifications predicted increased Go/No-Go commission errors and slower response times. Caffeine and mood had secondary roles.

---

9. References

- Apple Screen Time Documentation
- PsyToolkit.org – Go/No-Go Task
- Breiman, L. (2001). Random Forests. *Machine Learning*
- Field, A. (2013). *Discovering Statistics Using IBM SPSS Statistics*
- scikit-learn Documentation