

# ***THE COMPARISON OF MACHINE LEARNING MODEL FOR CREDIT CARD DEFAULT PREDICTION***



# LIST OF CONTENT



INTRODUCTION



OBJECTIVE



EXPLORATORY  
ANALYSIS



MODEL REPORT  
SUMMARIES



CONCLUSIONS

A close-up photograph of a person's hands typing on a silver laptop keyboard. The laptop is open, and the screen is visible. In the foreground, there is a smartphone and some papers. The background is blurred, showing what appears to be a person sitting at a desk.

# ***INTRODUCTION***

The use of **credit cards** today is increasingly widespread. One of the main reasons is **the ease of transactions**. However, banks as credit service providers often experience **problems related to losses** due to many **users who do not pay** off their **credit** on time. **Machine learning** can be used to **predict** potentially **default users** based on existing variables. So, it is hoped that it can be used as a **consideration** in order to **minimize the losses** that can be experienced.

A close-up photograph of a person's hands typing on a silver laptop keyboard. The laptop is open, and the screen is visible. In the foreground, there is a white smartphone and some papers. The background is blurred, showing a person sitting at a desk.

# ***OBJECTIVE***

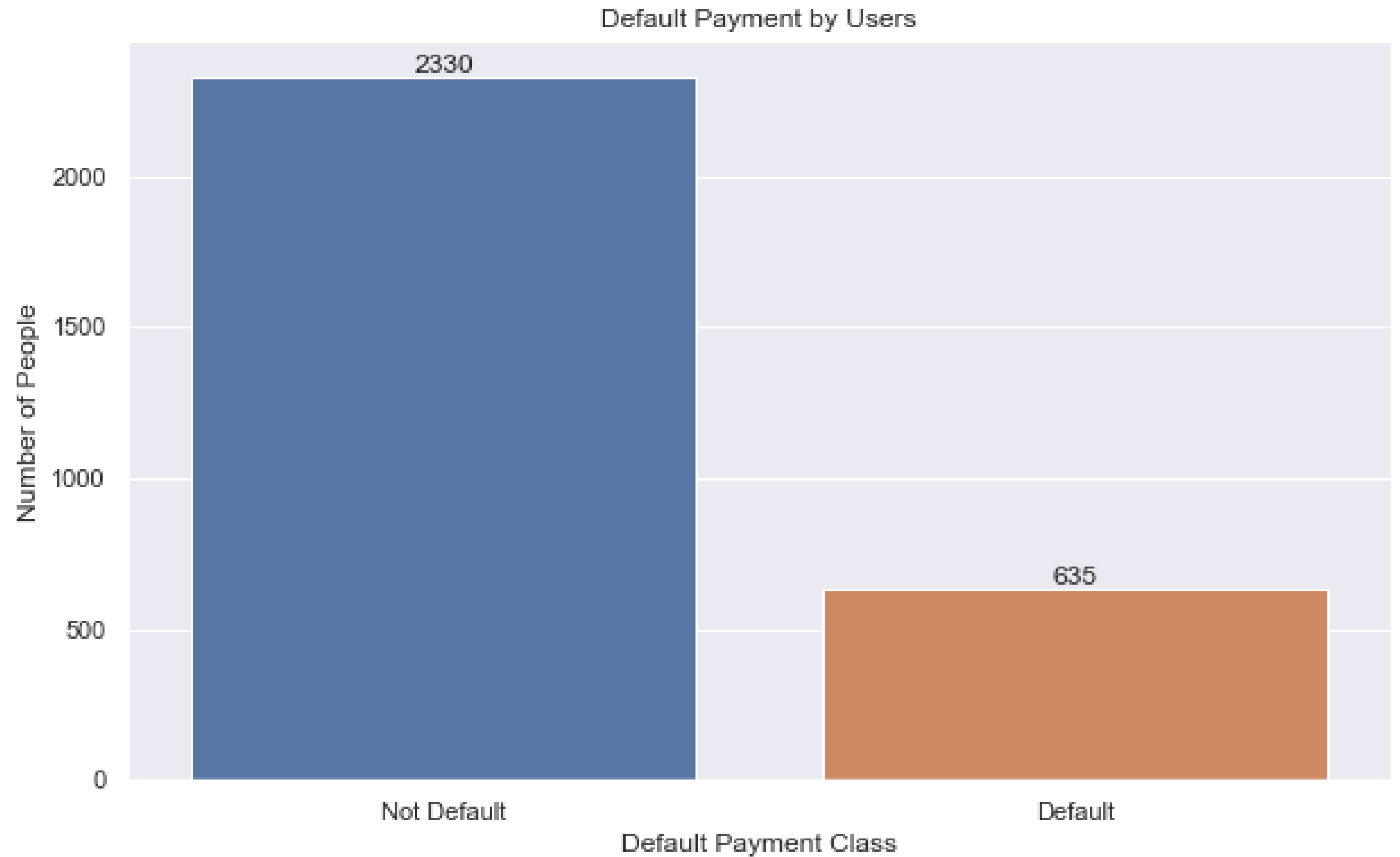
Create a machine learning **logistic regression** model, **support vector machine** (svm), **decision trees** model, **random forests** model, **k-nearest neighbors** (knn) model, **naive bayes** model, and **gradient boosting** model to predict default users on credit cards.

**Compare** machine learning models built to **predict user defaults** on credit cards and **determine the best model**.

Perform **hyperparameter tuning** for the **best prediction model**.

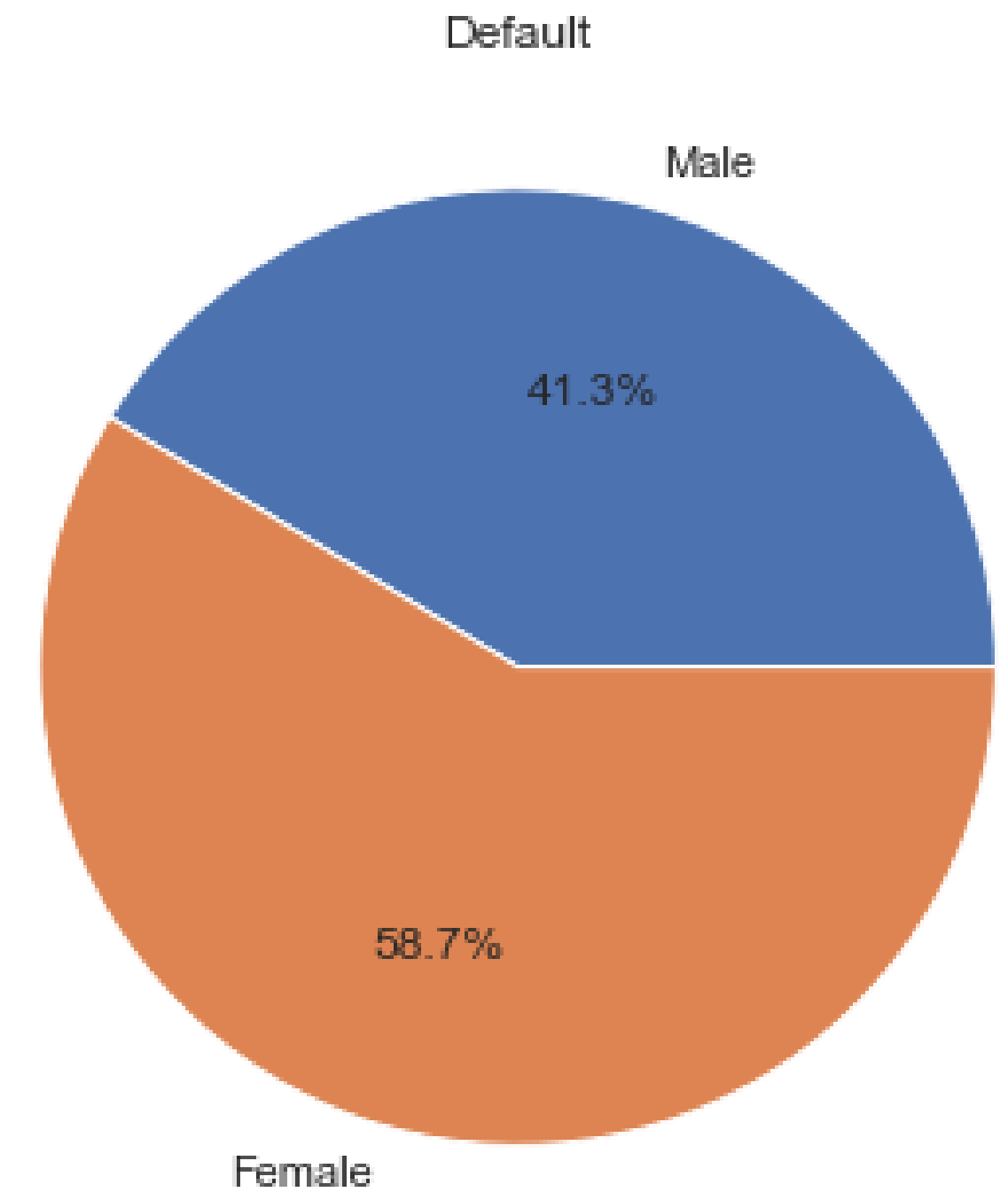
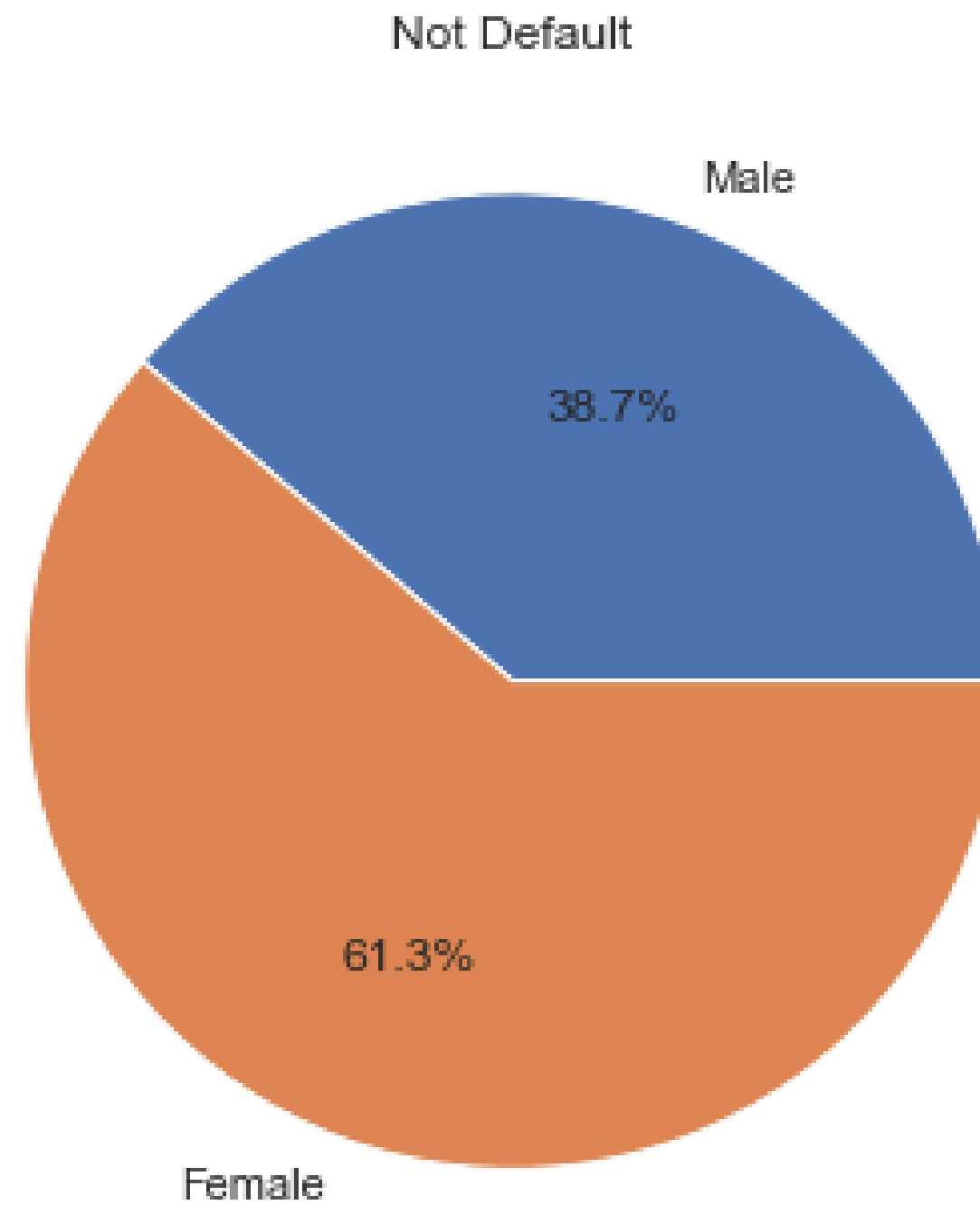
# ***EXPLORATORY ANALYSIS***

# Default Users



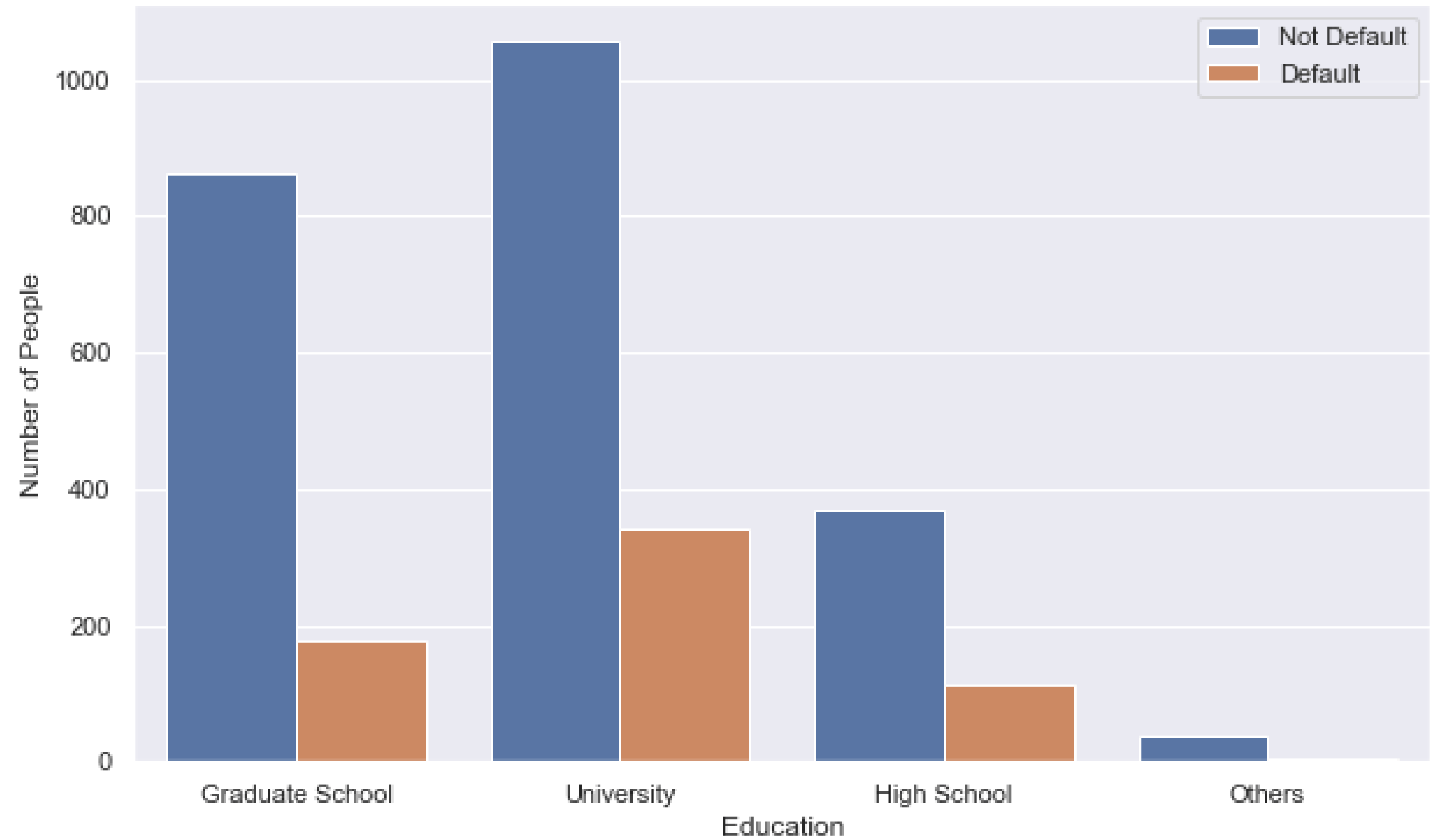
## User's Gender

### Gender Proportion on Default Payment



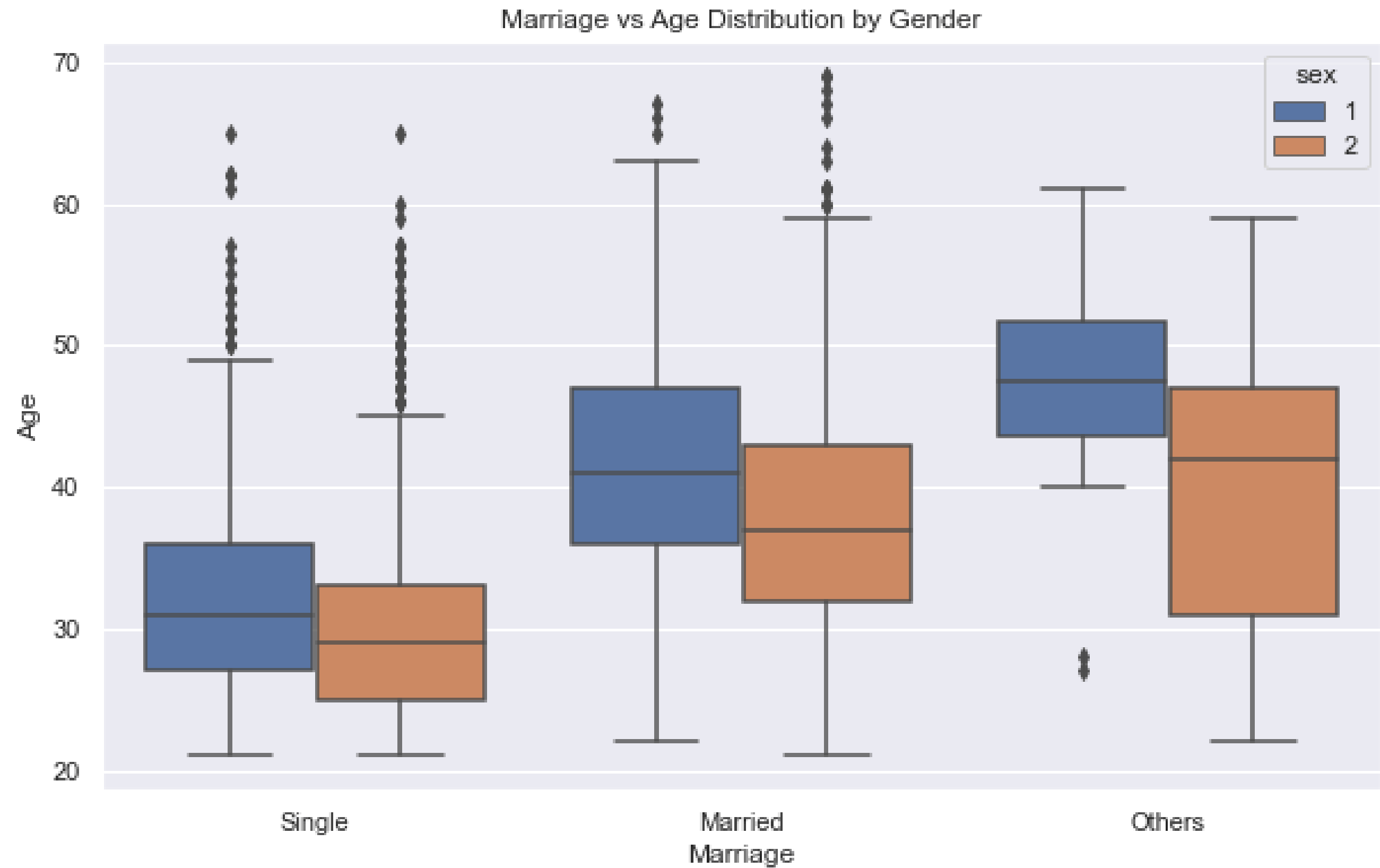
## ***User's Eductaion***

### Education Level

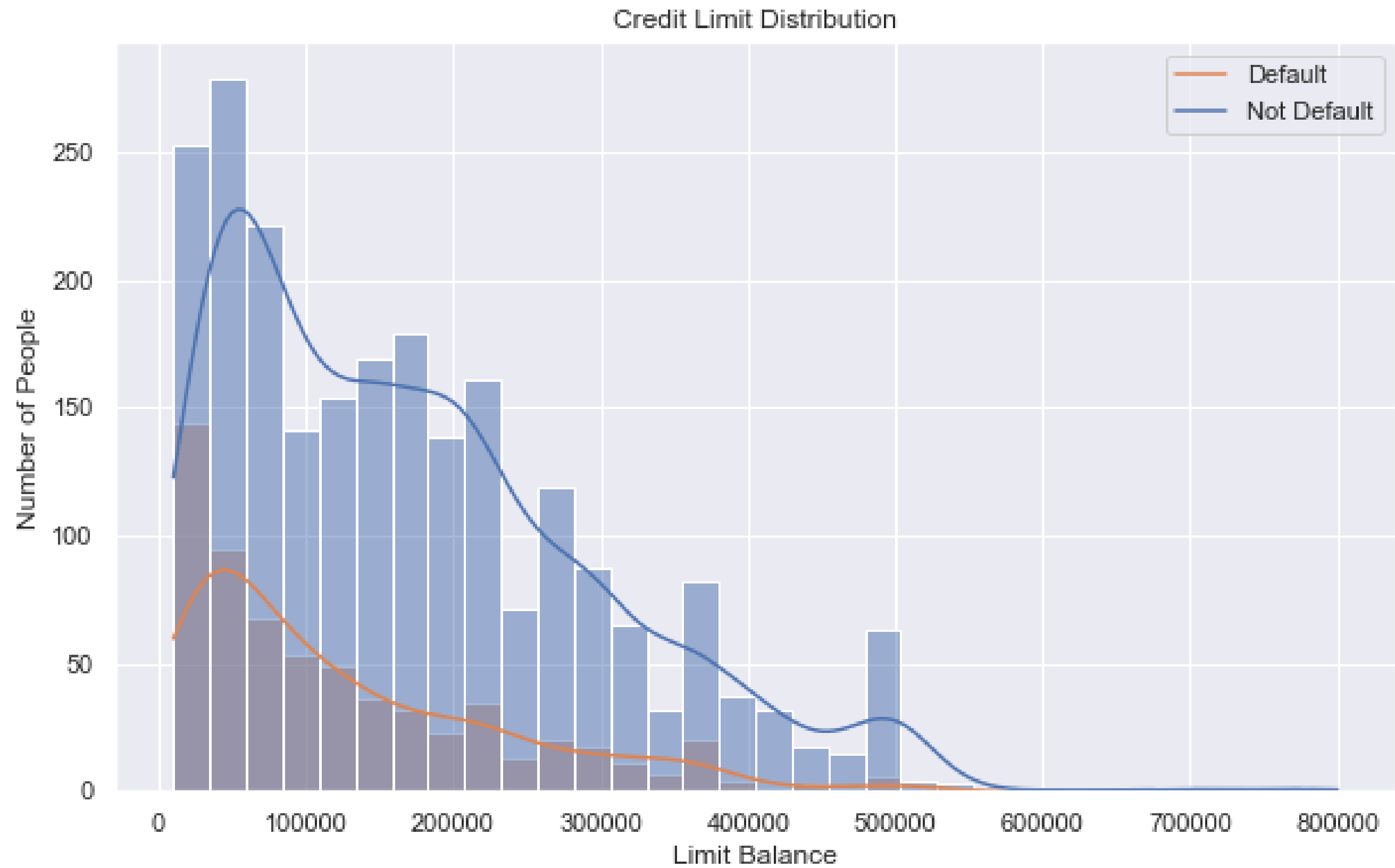




# Marriage Vs Age



# Credit Limit



# MODEL REPORT SUMMARIES

## Training Model

Metric	Logistic Regression	Support Vector Machine	Decision Tree	Random Forests	KNN	NAIVE BAYES	GRADIENT BOOSTING
Accuracy Score	0.83	0.84	0.98	0.98	0.85	0.81	0.84
Recall Label 1	0.38	0.44	0.91	0.93	0.50	0.44	0.39

## Testing Model

Metric	Logistic Regression	Support Vector Machine	Decision Tree	Random Forests	KNN	NAIVE BAYES	GRADIENT BOOSTING
Accuracy Score	0.86	0.86	0.77	0.83	0.84	0.82 (0.84)	0.86
Recall Label 1	0.43	0.49	0.47	0.46	0.47	0.51 (0.40)	0.45

# CONCLUSIONS



Some algorithms are not good in modeling (DT and RF) and experience overfitting caused by unbalanced data.



The best models in the modeling are KNN, SVM, and NB.



Hyperparameter tuning yang dilakukan pada model NB meningkatkan overall akurasi pada testing model NB, akan tetapi recall label 1 mengalami penurunan.