

Academic Success Dataset

```
import pandas as pd

from sklearn.model_selection import train_test_split
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import accuracy_score


# Load the data

train_df = pd.read_csv('train.csv')
test_df = pd.read_csv('test.csv')


# Explore the data

print(train_df.info())
print(train_df.describe())


# Handle missing values (if any)

train_df = train_df.fillna(train_df.mean())
test_df = test_df.fillna(test_df.mean())


# Encode categorical features

from sklearn.preprocessing import LabelEncoder

le = LabelEncoder()

train_df['Target'] = le.fit_transform(train_df['Target'])


# Split the data into train and validation sets

X_train, X_val, y_train, y_val = train_test_split(train_df.drop('Target', axis=1), train_df['Target'],
test_size=0.2, random_state=42)


# Train a logistic regression model

model = LogisticRegression()

model.fit(X_train, y_train)
```

```
# Evaluate the model on the validation set
```

```
y_pred = model.predict(X_val)
```

```
val_accuracy = accuracy_score(y_val, y_pred)
```

```
print(f'Validation Accuracy: {val_accuracy:.2f}')
```

```
# Make predictions on the test set
```

```
test_predictions = model.predict(test_df)
```

```
# Create the submission file
```

```
submission = pd.DataFrame({'id': test_df.index, 'Target': le.inverse_transform(test_predictions)})
```

```
submission.to_csv('submission.csv', index=False)
```