In [1]:
```
#Razat Siwakoti (A00046635)
#DMV302 - Assessment 2
#NN1.ipynb created on Jupyter notebook


#source: Scikit-learn(2015)
#https://scikit-learn.org/stable/modules/generated/sklearn.neural_network.MLPClassifier.

#Deepika S. (2019)
#https://www.pluralsight.com/guides/machine-learning-neural-networks-scikit-learn
```

In [18]:
```
#importing necessary libraries
import pandas as pd
from sklearn.neural_network import MLPClassifier
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score, confusion_matrix, mean_squared_error
import matplotlib.pyplot as plt
import seaborn as sns
import numpy as np
```

In [19]:
```
# Load the dataset
df = pd.read_csv("AtRiskStudentsTraining.csv")
df.head()
```

Out[19]:

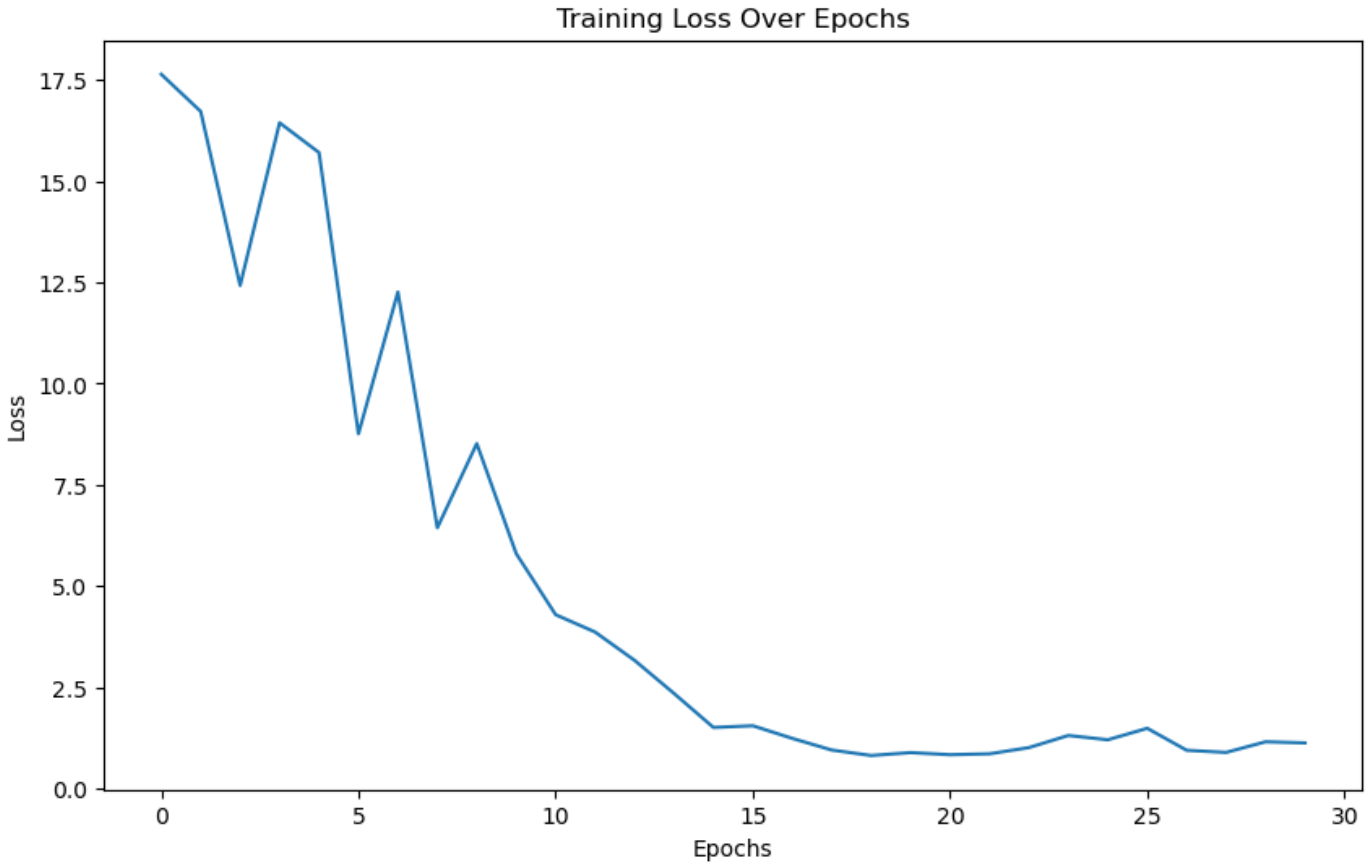
	GPA	attendance	duration	language	at-risk
0	2.07	76	9586	41	1
1	1.97	19	3772	28	0
2	2.49	43	1506	10	1
3	1.94	82	9223	3	1
4	0.52	37	8232	64	0

In [20]:
```
# Separate features and target variable
X = df.drop('at-risk', axis=1)
y = df['at-risk']

# Split the dataset into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42
```

In [21]:
```
# Define and train the neural network
activation_function = 'relu'
num_neurons = 50

# Create the neural network model
# Implementation of activation function and number of neurons in the hidden layer
model = MLPClassifier(hidden_layer_sizes=(num_neurons,), activation=activation_function,
```

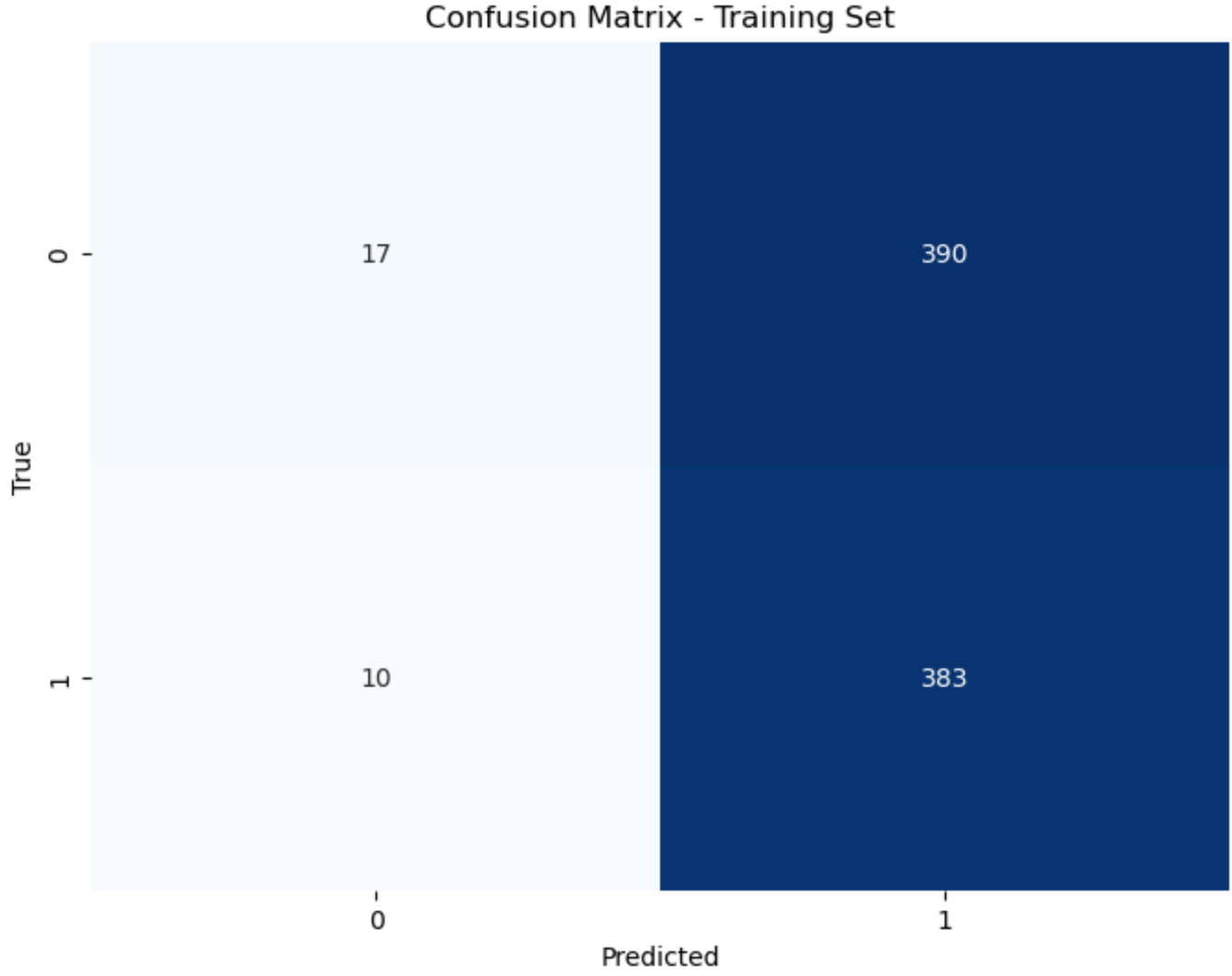
In [22]:
```
# Train the model
history = model.fit(X_train, y_train)
```

In [23]:
```
# Visualize training loss over epochs
plt.figure(figsize=(10, 6))
plt.plot(history.loss_curve_)
plt.title('Training Loss Over Epochs')
plt.xlabel('Epochs')
plt.ylabel('Loss')
plt.show()
```



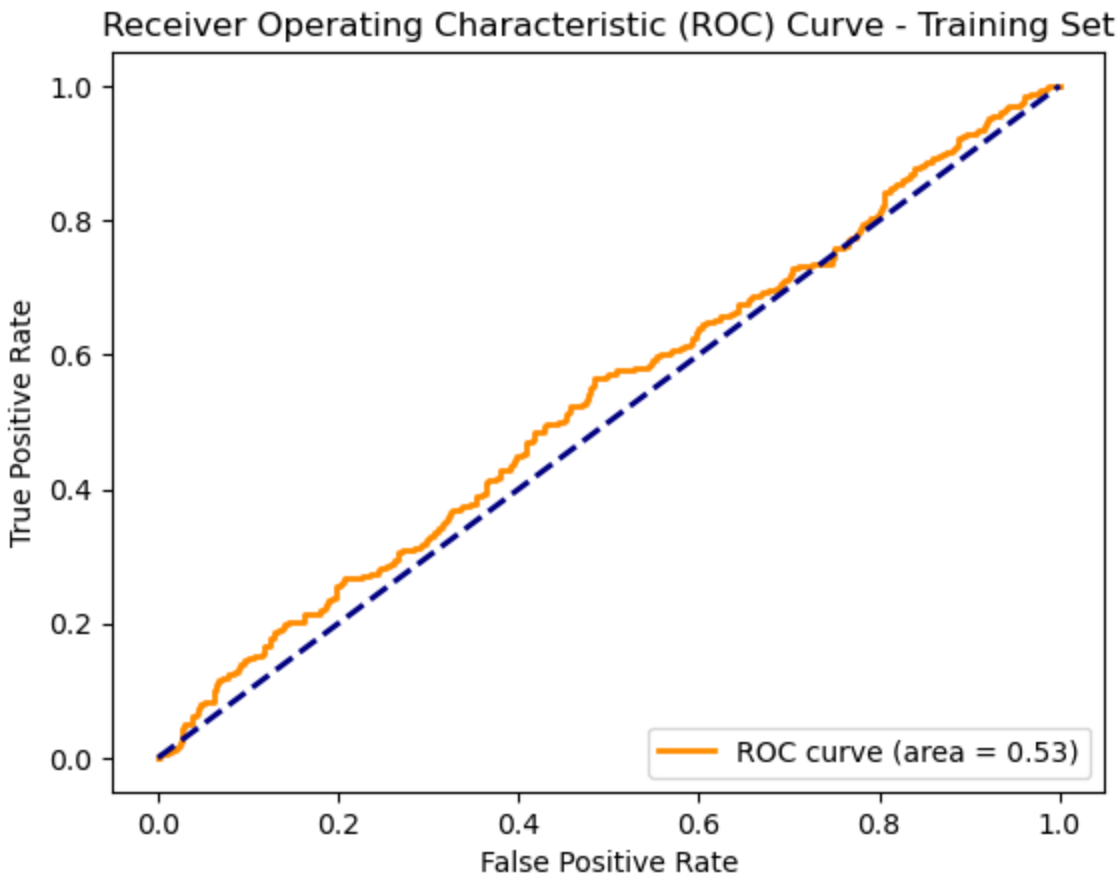
In [24]:
```
# Predictions on the training set for demonstration purposes
train_predictions = model.predict(X_train)

# Evaluate the model on the training set
train_accuracy = accuracy_score(y_train, train_predictions)
print(f"Training Accuracy: {train_accuracy:.2f}")

#Confusion Matrix
# Visualize confusion matrix
plt.figure(figsize=(8, 6))
sns.heatmap(confusion_matrix(y_train, train_predictions), annot=True, fmt='d', cmap='Blu
plt.title('Confusion Matrix - Training Set')
plt.xlabel('Predicted')
plt.ylabel('True')
plt.show()
```

```
Training Accuracy: 0.50
```

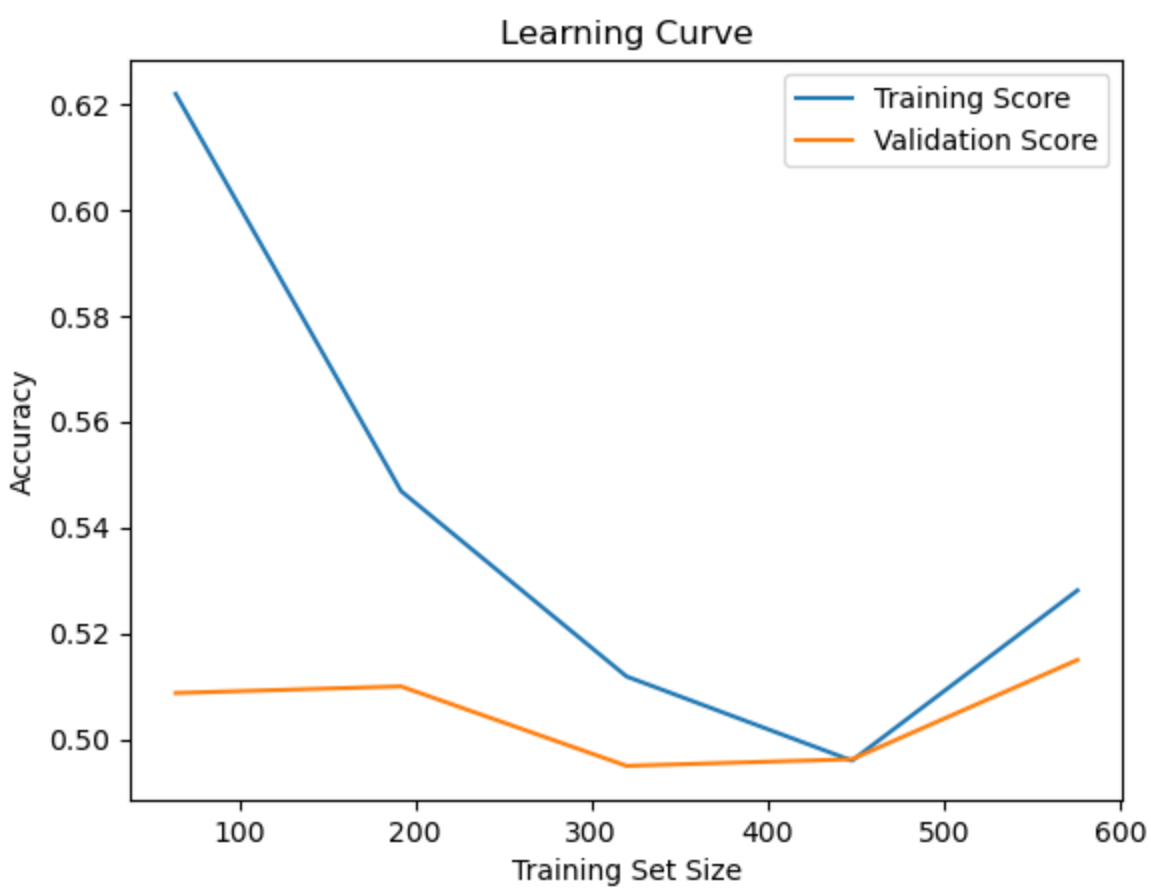


In [25]:

```
#Roc curve
from sklearn.metrics import roc_curve, auc
# Calculate the ROC curve and area under the curve (AUC)
fpr, tpr, thresholds = roc_curve(y_train, model.predict_proba(X_train)[:, 1])
roc_auc = auc(fpr, tpr)
# Plot the ROC curve
plt.plot(fpr, tpr, color='darkorange', lw=2, label='ROC curve (area = {:.2f})'.format(ro
# Plot the diagonal line representing random guessing
plt.plot([0, 1], [0, 1], color='navy', lw=2, linestyle='--')
# Add labels and title to the plot
plt.xlabel('False Positive Rate')
plt.ylabel('True Positive Rate')
plt.title('Receiver Operating Characteristic (ROC) Curve - Training Set')
# Add a legend to the plot
plt.legend(loc='lower right')
plt.show()
```



In [26]:
```
#Learning Curve
from sklearn.model_selection import learning_curve
#Calculate Learning curve
train_sizes, train_scores, valid_scores = learning_curve(model, X_train, y_train, train_
# Plot the learning curve and add labels to it
plt.plot(train_sizes, np.mean(train_scores, axis=1), label='Training Score')
plt.plot(train_sizes, np.mean(valid_scores, axis=1), label='Validation Score')
plt.xlabel('Training Set Size')
plt.ylabel('Accuracy')
plt.title('Learning Curve')
plt.legend()
plt.show()
```


Learning Curve
Training Score
Validation Score
Accuracy
0.62
0.60
0.58
0.56
0.54
0.52
0.50
100
200
300
400
500
600
Training Set Size

In [ ]: