

Model Development Phase Template

Date	23 september 2024
Team ID	LTVIP2024TMID25030
Project Title	FAKE NEWS ANALYSIS IN SOCIAL MEDIA
Maximum Marks	4 Marks

Initial Model Training Code, Model Validation and Evaluation Report

The initial model training code will be showcased in the future through a screenshot. The model validation and evaluation report will include classification reports, accuracy, and confusion matrices for multiple models, presented through respective screenshots.

Initial Model Training Code:

```
# Initialize and train the model
model = MultinomialNB()
model.fit(X_train, y_train)

# Make predictions
y_pred = model.predict(X_test)

# Accuracy Score
accuracy = accuracy_score(y_test, y_pred)
print(f'Accuracy: {accuracy * 100:.2f}%')

# Classification Report
print(classification_report(y_test, y_pred))

# Confusion Matrix
cm = confusion_matrix(y_test, y_pred)
sns.heatmap(cm, annot=True, fmt='d', cmap='Blues')
plt.title('Confusion Matrix')
plt.show()
```

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

```
# Initialize CountVectorizer
vectorizer = CountVectorizer(stop_words='english')

# Fit and transform the cleaned text data
X = vectorizer.fit_transform(df['cleaned_text'])

# Converting the labels into binary format (1 for FAKE, 0 for REAL)
df['label_num'] = df['label'].apply(lambda x: 1 if x == 'FAKE' else 0)
y = df['label_num']
```

Model Validation and Evaluation Report:

Model	Classification Report	Accuracy	Confusion Matrix																																				
Multinomial Naïve Bayes	<pre>print(classification_report(y_test, y_pred))</pre> <table><thead><tr><th></th><th>precision</th><th>recall</th><th>f1-score</th><th>support</th></tr></thead><tbody><tr><td>0</td><td>0.92</td><td>0.88</td><td>0.90</td><td>1247</td></tr><tr><td>1</td><td>0.78</td><td>0.85</td><td>0.81</td><td>610</td></tr><tr><td>accuracy</td><td></td><td></td><td>0.87</td><td>1857</td></tr><tr><td>macro avg</td><td>0.85</td><td>0.87</td><td>0.86</td><td>1857</td></tr><tr><td>weighted avg</td><td>0.88</td><td>0.87</td><td>0.87</td><td>1857</td></tr></tbody></table>		precision	recall	f1-score	support	0	0.92	0.88	0.90	1247	1	0.78	0.85	0.81	610	accuracy			0.87	1857	macro avg	0.85	0.87	0.86	1857	weighted avg	0.88	0.87	0.87	1857	87.18%	<pre>cm = confusion_matrix(y_test, y_pred)</pre> <table><thead><tr><th colspan="2">Confusion Matrix</th></tr></thead><tbody><tr><td>[[1103</td><td>144]</td></tr><tr><td>[94</td><td>516]]</td></tr></tbody></table>	Confusion Matrix		[[1103	144]	[94	516]]
	precision	recall	f1-score	support																																			
0	0.92	0.88	0.90	1247																																			
1	0.78	0.85	0.81	610																																			
accuracy			0.87	1857																																			
macro avg	0.85	0.87	0.86	1857																																			
weighted avg	0.88	0.87	0.87	1857																																			
Confusion Matrix																																							
[[1103	144]																																						
[94	516]]																																						