

Assignment 2

Team members: Benedetta Pacilli, Valentina Pieri, Parisa Sokuti

Industry 4.0 - Assignment 2

We experimented with two other classifiers: Random Forest and AdaBoost.

Random Forest code

```
from sklearn.ensemble import RandomForestClassifier

rf = RandomForestClassifier(n_estimators=200, max_depth=2,
random_state=0).fit(feature_train, label_train)

rf_pred = rf.predict(feature_test)

accuracy_rf = rf.score(feature_test, label_test)
print("Accuracy Random Forest: ", accuracy_rf)

cm_rf = confusion_matrix(label_test, rf_pred)
print("CM Random Forest")
print(cm_rf)

ConfusionMatrixDisplay.from_estimator(rf, feature_test, label_test)

print("New Feature:")
new_feature = feature_test.tail(2)
display(new_feature)
new_rf_pred_2 = rf.predict(new_feature)
new_rf_proba = rf.predict_proba(new_feature)
print("New Random Forest Prediction:")
display(new_rf_pred)
display(new_rf_proba)
```

AdaBoost code

```
from sklearn.ensemble import AdaBoostClassifier

ada = AdaBoostClassifier(n_estimators=200,
random_state=0).fit(feature_train, label_train)

ada_pred = ada.predict(feature_test)

accuracy_ada = ada.score(feature_test, label_test)
```

```

print("Accuracy AdaBoost: ", accuracy_ada)

cm_ada = confusion_matrix(label_test, ada_pred)
print("CM AdaBoost")
print(cm_ada)

ConfusionMatrixDisplay.from_estimator(ada, feature_test, label_test)

print("New Feature:")
new_feature = feature_test.tail(2)
display(new_feature)
new_ada_pred_2 = ada.predict(new_feature)
new_ada_pred = ada.predict_proba(new_feature)
print("New AdaBoost Prediction:")
display(new_ada_pred)
display(new_ada_pred_2)

```

Result analysis between Linear SVM and our models

Accuracies and Confusion Matrixes

```

Accuracy SVM Linear Kernel: 0.7792207792207793
Accuracy Random Forest: 0.8441558441558441
Accuracy AdaBoost: 0.8961038961038961

```

CM SVM Linear:

```

[[22 0 4]
 [ 2 21 3]
 [ 7 1 17]]

```

CM Random Forest:

```

[[22 0 4]
 [ 1 25 0]
 [ 6 1 18]]

```

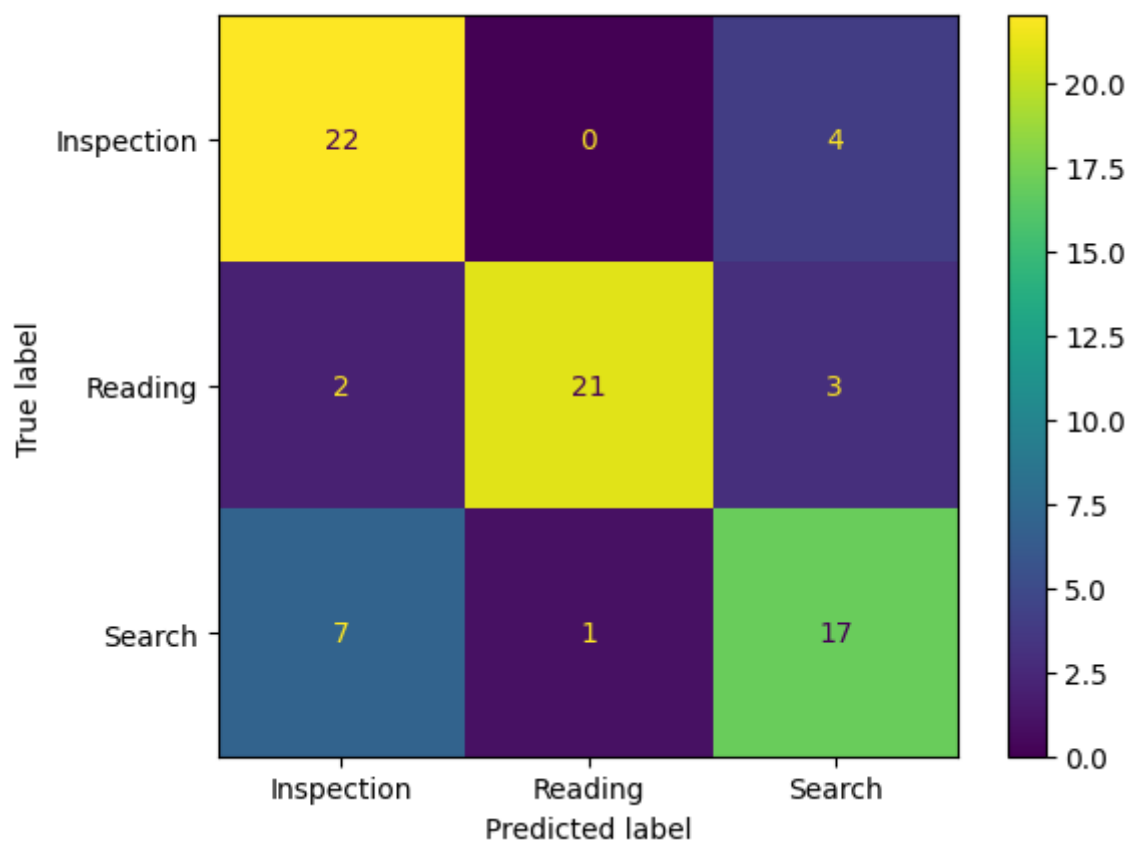
CM AdaBoost:

```

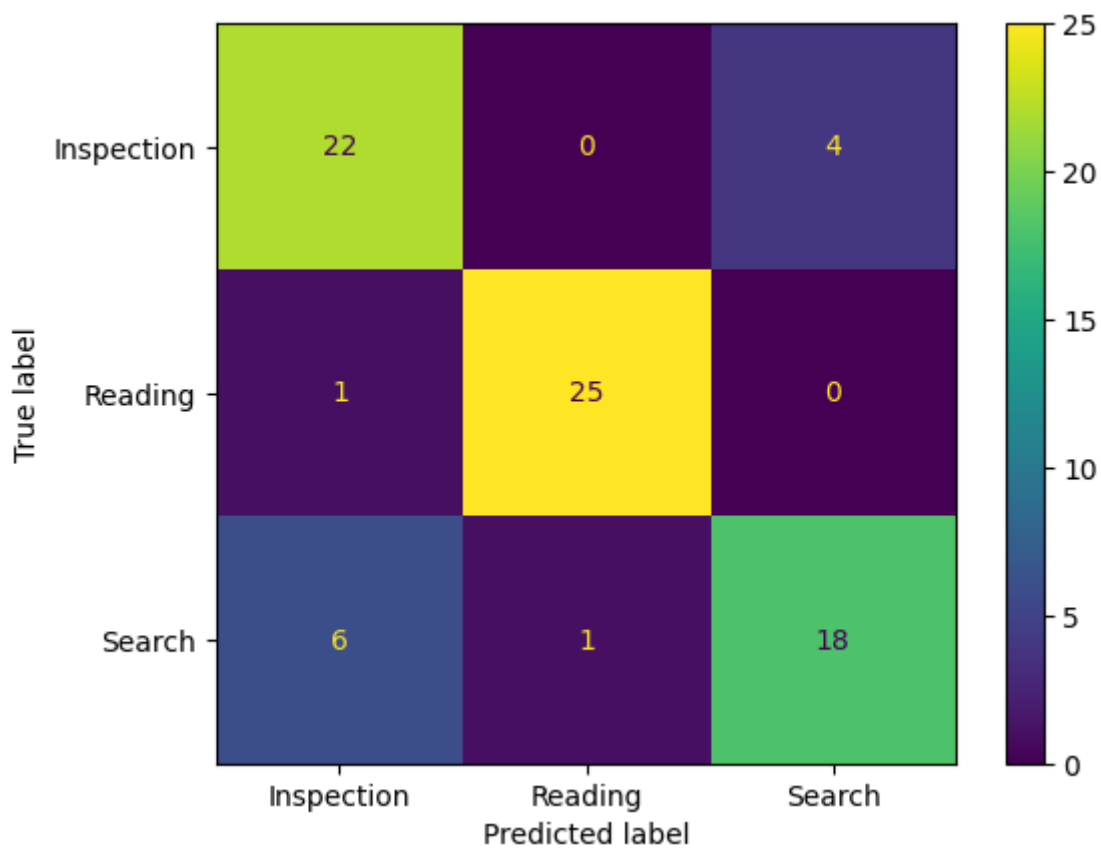
[[22 0 4]
 [ 1 25 0]
 [ 3 0 22]]

```

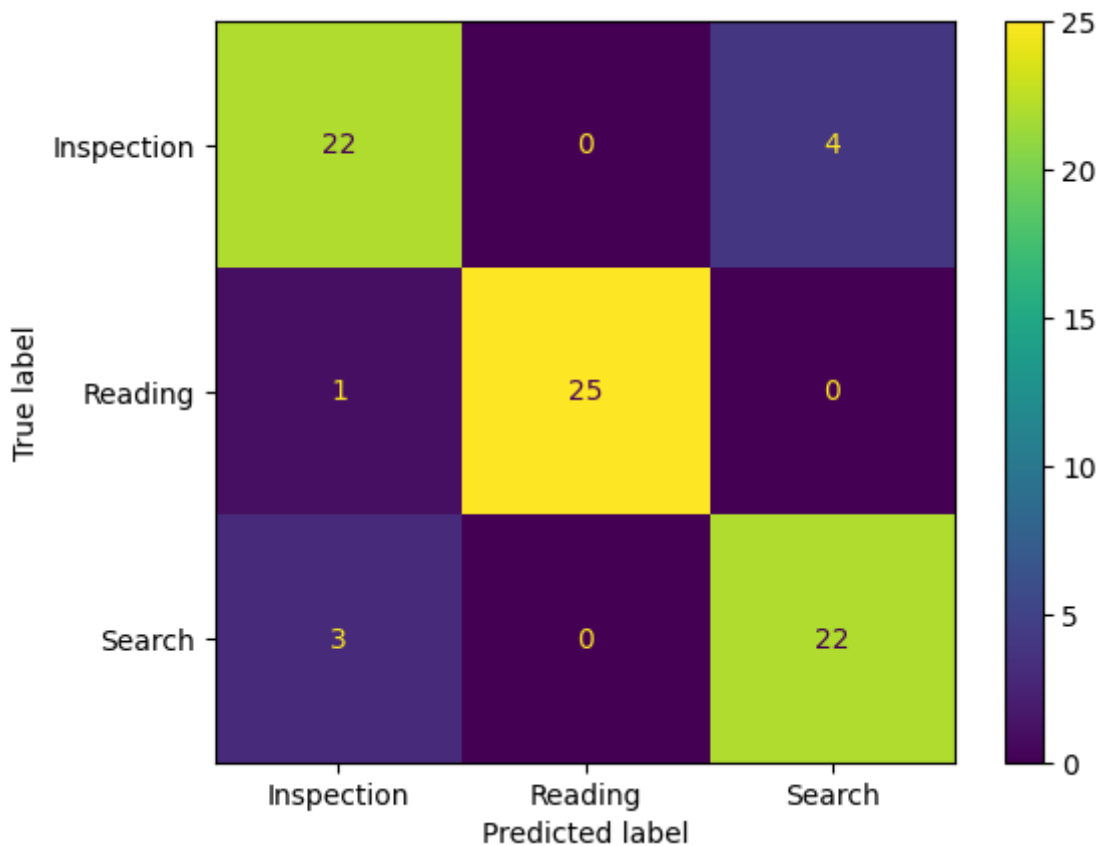
Confusion Matrix SVM:



Confusion Matrix RandomForest:



Confusion Matrix AdaBoost:



New feature prediction results

New Linear Prediction:

```
array([[0.08657579, 0.53616673, 0.37725748], [0.10471277, 0.8188641 ,
0.07642312]])
array(['Reading', 'Reading'], dtype=object)
```

New Random Forest Prediction:

```
array([[0.07131885, 0.84797325, 0.0807079 ], [0.08901043, 0.84645339,
0.06453618]])
array(['Reading', 'Reading'], dtype=object)
```

New AdaBoost Prediction:

```
array([[0.21957406, 0.4564633 , 0.32396263], [0.16951591, 0.55743797,
0.27304612]])
array(['Reading', 'Reading'], dtype=object)
```

Conclusions on the results

As shown from the results, the models that we used perform better than SVM Linear Kernel model.

AdaBoost and RandomForest are ensemble methods, which often perform well when the relationships between features and the target variable are nonlinear or complex. Random Forest and AdaBoost can capture intricate patterns in the data through the combination of multiple weak learners, whereas SVM with a linear kernel might struggle to capture such complexity.

Moreover, ensemble methods are typically robust to noisy data and outliers. Random Forest, for example, builds multiple decision trees and averages their predictions, which can reduce the impact of noise. AdaBoost, by focusing on misclassified instances in successive iterations, can also adapt well to noisy data.

What are fixations and saccades?

Fixations are periods when the eyes are relatively still, focused on a certain point.

Saccades are fast and simultaneous movements of the eyes, between two phases of fixations, in the same direction.

For example, when a person is reading, when the eyes are looking at a word it is called a fixation. When the reader goes from one word to another it is called a saccade.