Worked example

02 July 2024 12:18

Example Scenario: Fraud Detection

Imagine we have a dataset of credit card transactions where only 1% of the transact
(positive class), and the remaining 99% are legitimate transactions (negative class).

Dataset Characteristics:

- Total transactions: 10,000
- Fraudulent transactions (positive class): 100
- Legitimate transactions (negative class): 9,900

Model Performance Evaluation

Let's consider a machine learning model trained to classify these transactions as frau
legitimate. After training, the model is evaluated using a confusion matrix, which bre
predictions as follows:

- **True Positives (TP)**: Predicted as fraudulent and actually fraudulent.
- **True Negatives (TN)**: Predicted as legitimate and actually legitimate.
- **False Positives (FP)**: Predicted as fraudulent but actually legitimate (Type I err
- **False Negatives (FN)**: Predicted as legitimate but actually fraudulent (Type II e

Assume the model's predictions are as follows:

- TP = 80 (correctly identified fraudulent transactions)
- TN = 9,800 (correctly identified legitimate transactions)
- FP = 100 (legitimate transactions incorrectly identified as fraudulent)
- FN = 20 (fraudulent transactions incorrectly identified as legitimate)

Confusion Matrix:

	Predicted Fraudulent	**Predicted Legitimate**
Actual Fraudulent	80 (TP)	20 (FN)
Actual Legitimate	100 (FP)	9,800 (TN)

tions are fraudulent

udulent or
eaks down

or).
error).

Calculating Metrics

Accuracy:

Accuracy=TP+TNTP+TN+FP+FN=80+9,80080+9,800+100+20=9,88010,000=0.988\tex
\frac{TP + TN}{TP + TN + FP + FN} = \frac{80 + 9,800}{80 + 9,800 + 100 + 20} = \frac{9
0.988Accuracy=TP+TN+FP+FNTP+TN=80+9,800+100+2080+9,800=10,0009,880=0.98

Recall (Sensitivity):

Recall=TPTP+FN=8080+20=80100=0.8\text{Recall} = \frac{TP}{TP + FN} = \frac{80}{8
{100} = 0.8Recall=TP+FNTP=80+2080=10080=0.8 Recall is 80%.

Interpretation

- **Accuracy**: The model shows high accuracy (98.8%), which might initially sugge performance. However, this high accuracy is mainly driven by the large numbe identified legitimate transactions (TN). It does not reflect the model's perform fraudulent transactions effectively.
- **Recall**: The recall score is 80%, indicating that the model correctly identifies 8 transactions. This metric is crucial in fraud detection because missing fraudule (false negatives) can be costly. A recall score of 80% means that the model cat portion of fraudulent activities, which is often more important than overall ac

Conclusion

In the context of imbalanced classes like fraud detection, where the positive class (f transactions) is rare compared to the negative class (legitimate transactions), using provides a more meaningful evaluation of the model's effectiveness. It directly mea ability to detect instances of the minority class (fraudulent transactions) accurately, decision-making in such applications. Therefore, despite high accuracy, the focus on that the model performs well where it matters most — identifying fraudulent activit financial losses.

t{Accuracy} =
),880}{10,000} =
38 Accuracy is 98.8%.

0 + 20} = \frac{80}

st good
er of correctly
nance on identifying

0% of all fraudulent
ent transactions
tches a significant
curacy.

raudulent
recall as a metric
sures the model's
which is critical for
recall helps ensure
ties to prevent