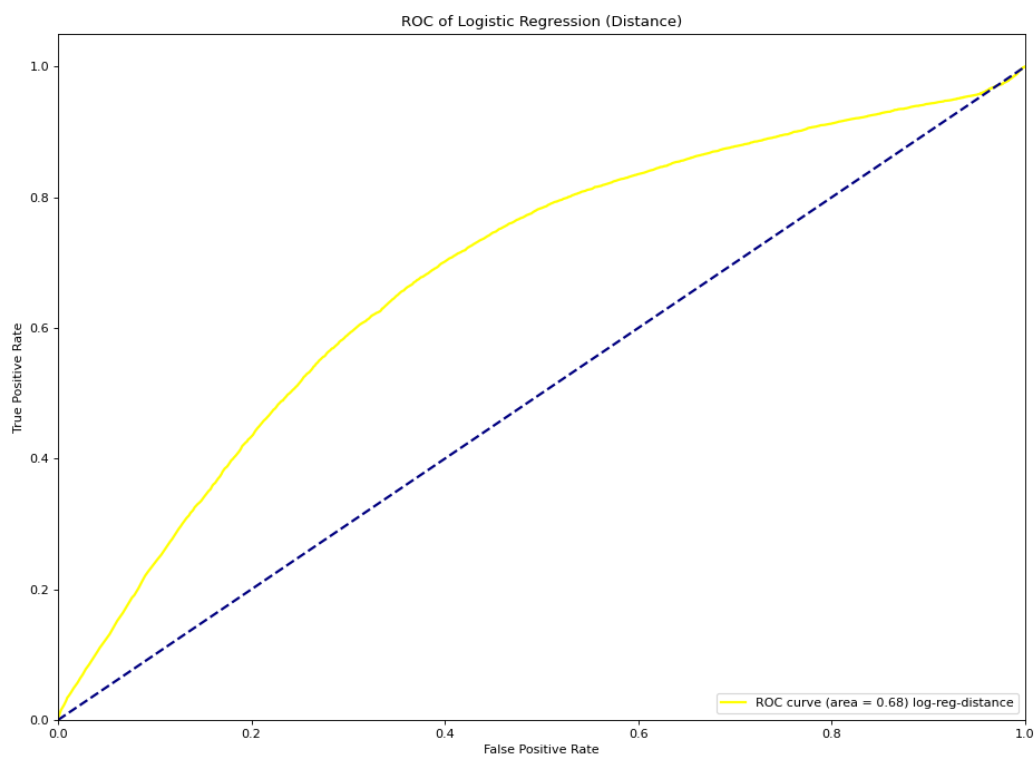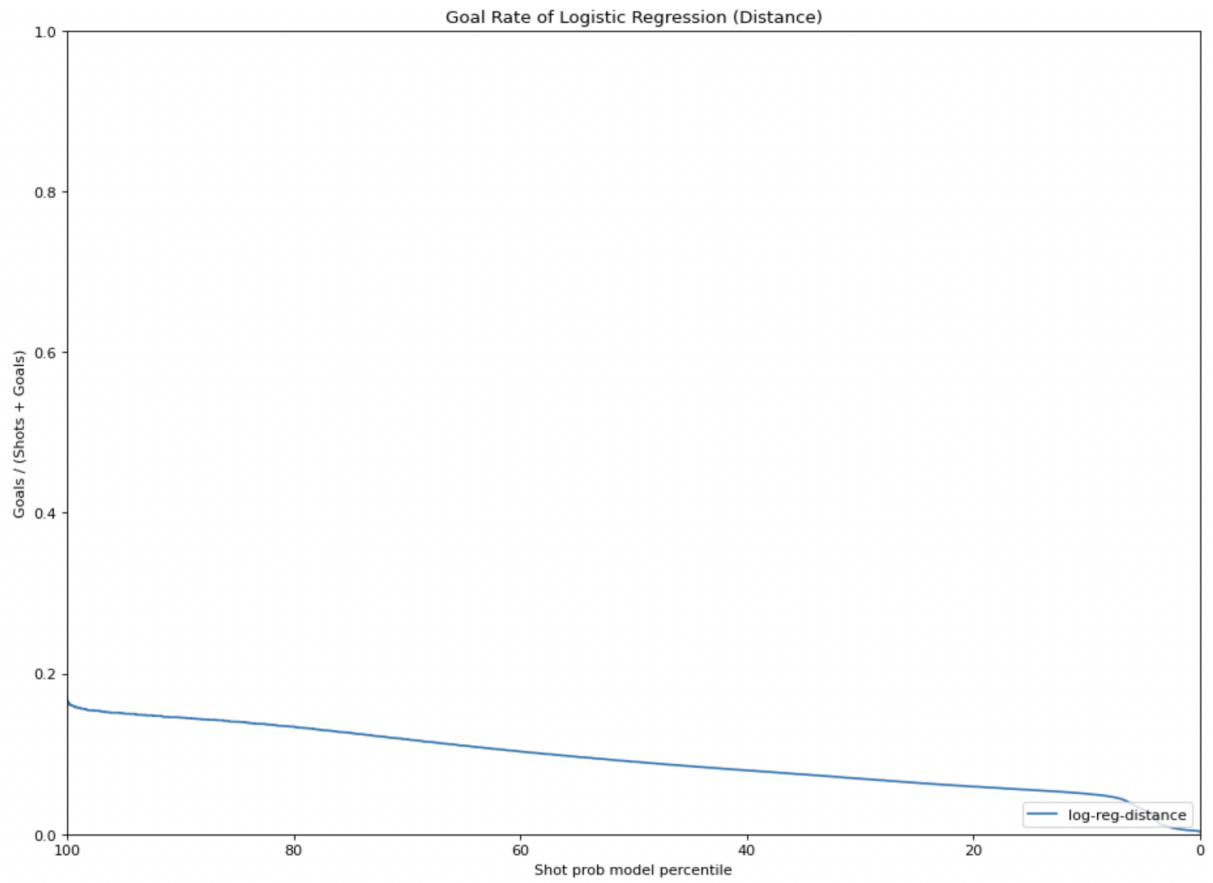## BASELINE MODELS

We use simple models like logistic regression to perform classification. We generated following models:
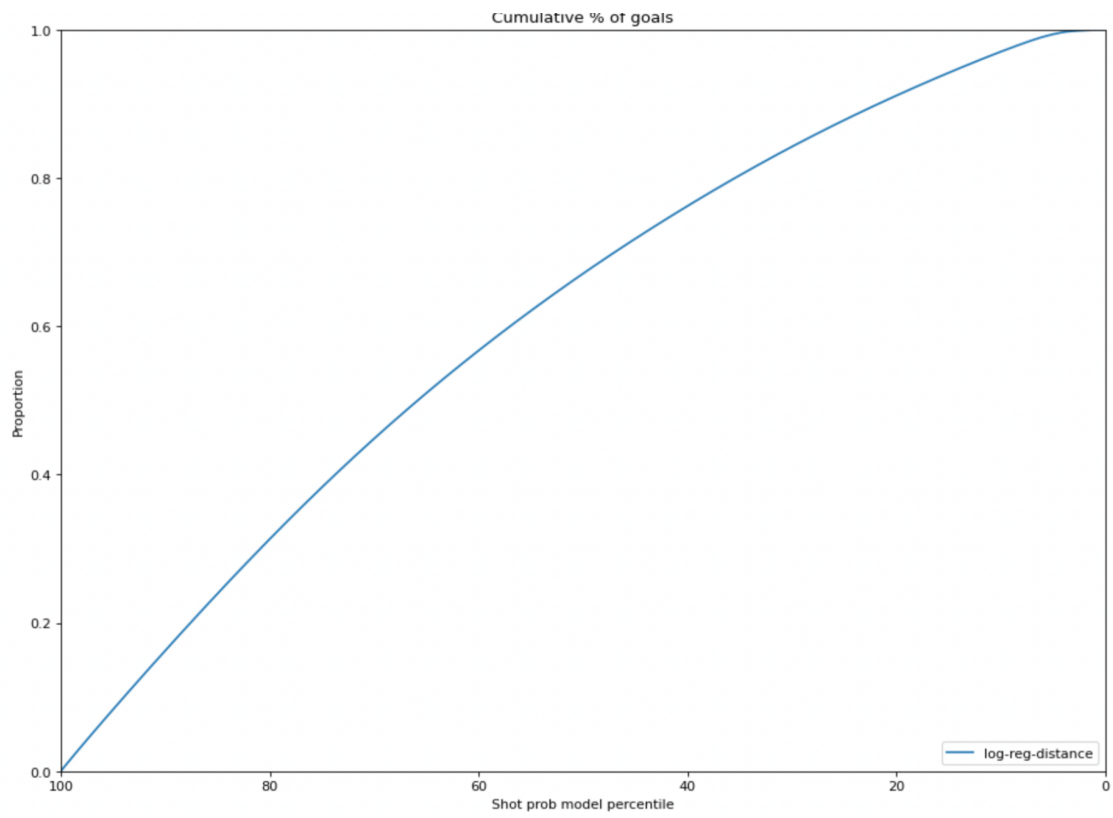
Using distance as feature:
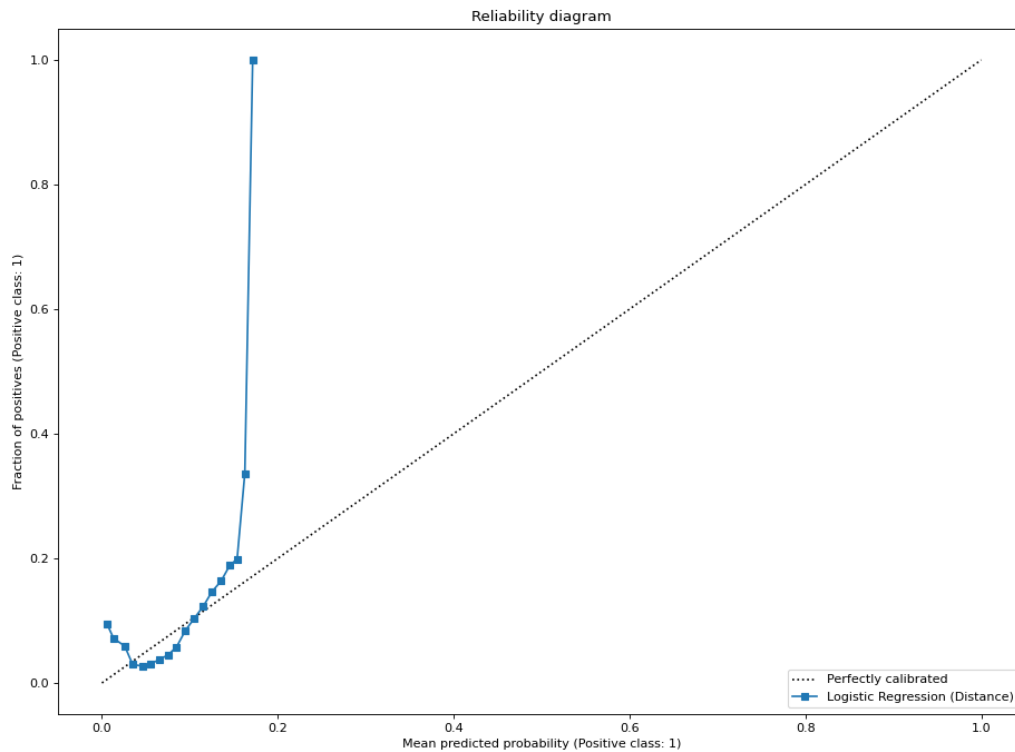We applied logistic regression with input feature as distance and output/label as is_goal.
We got the accuracy of 91% precision score 0.91 and Fscore 0.95. This means it can predict 91 out of 100 inputs correctly.

ROC of Logistic Regression (Distance)

Goal Rate of Logistic Regression (Distance)
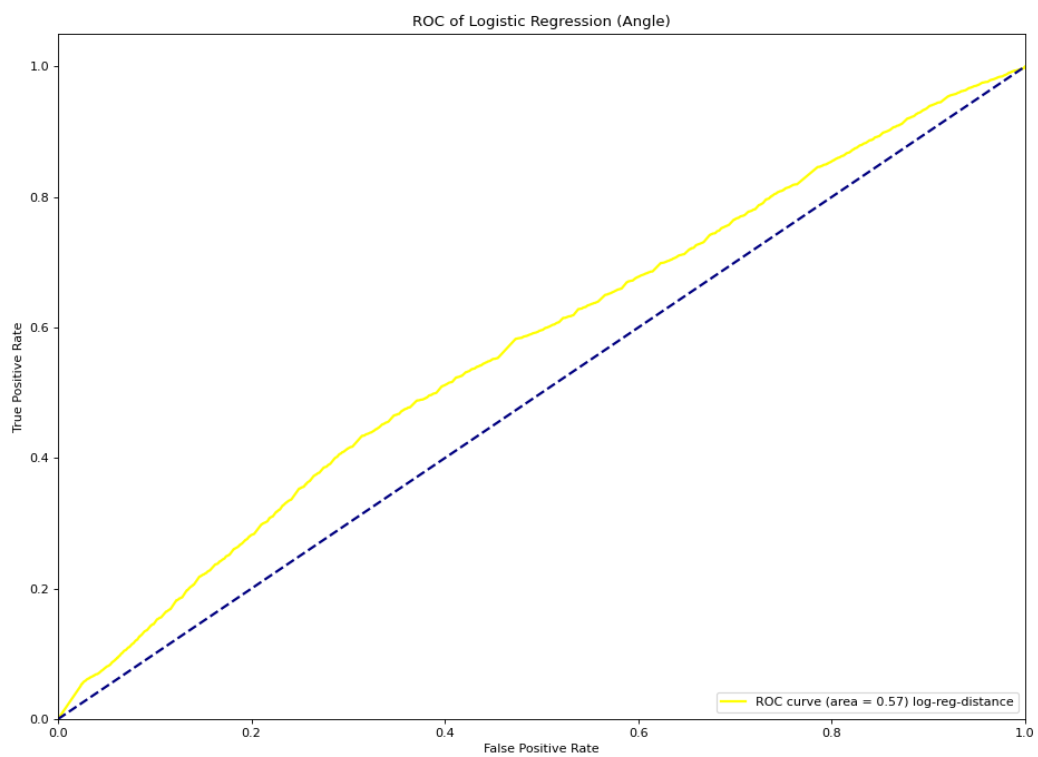
Cumulative % of goals
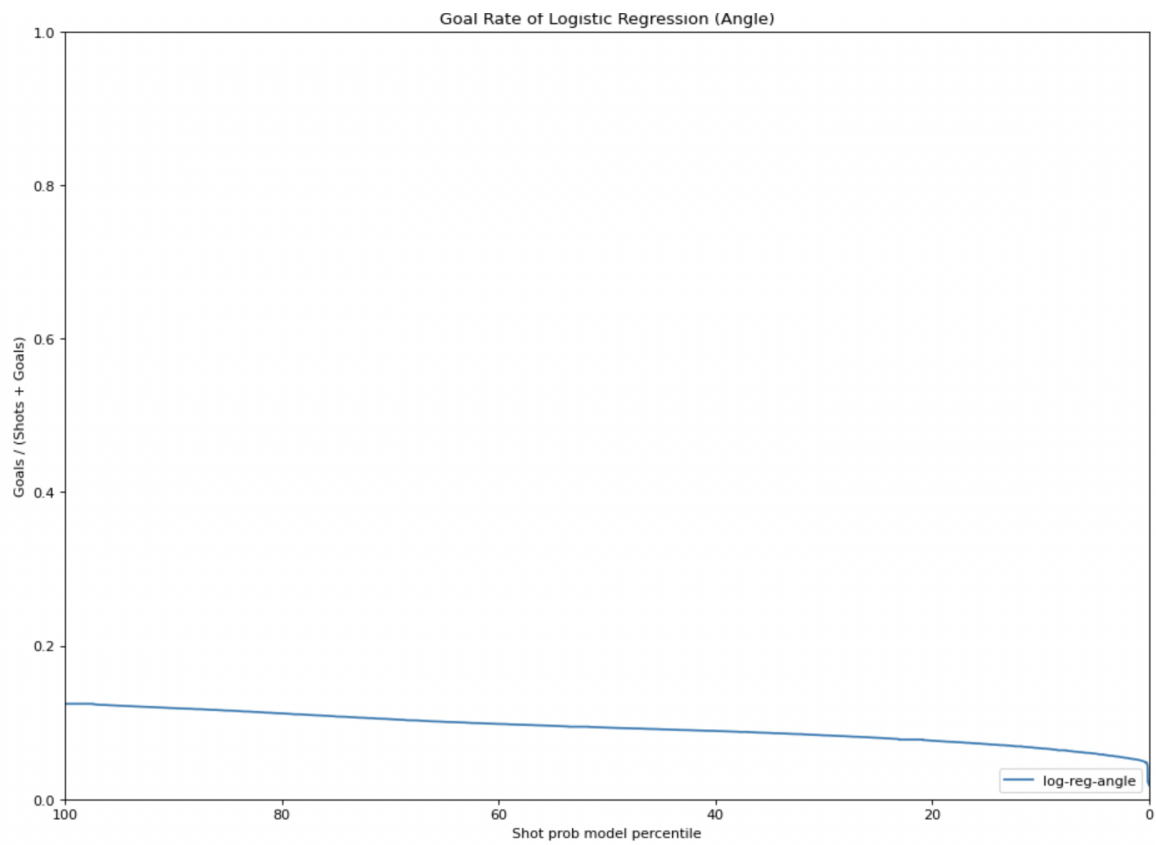
Reliability diagram

We applied logistic regression with input feature as angle and output/label as is_goal.
We got the accuracy of 91% precision score 0.91 and Fscore 0.95. This means it can predict 91 out of 100 inputs correctly.
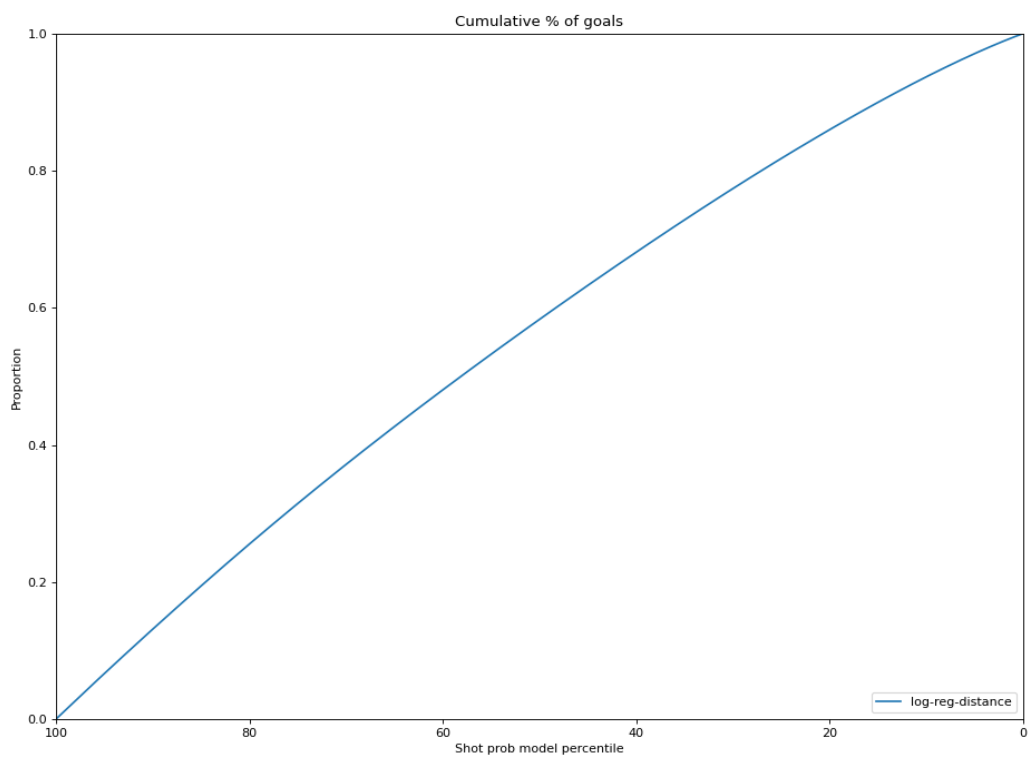
ROC of Logistic Regression (Angle)

Goal Rate of Logistic Regression (Angle)

Cumulative % of goals

Reliability diagram

We applied logistic regression with input feature as distance and output/label as is_goal.
We got the accuracy of 91% precision score 0.91 and Fscore 0.95. This means it can predict 91 out of 100 inputs correctly.
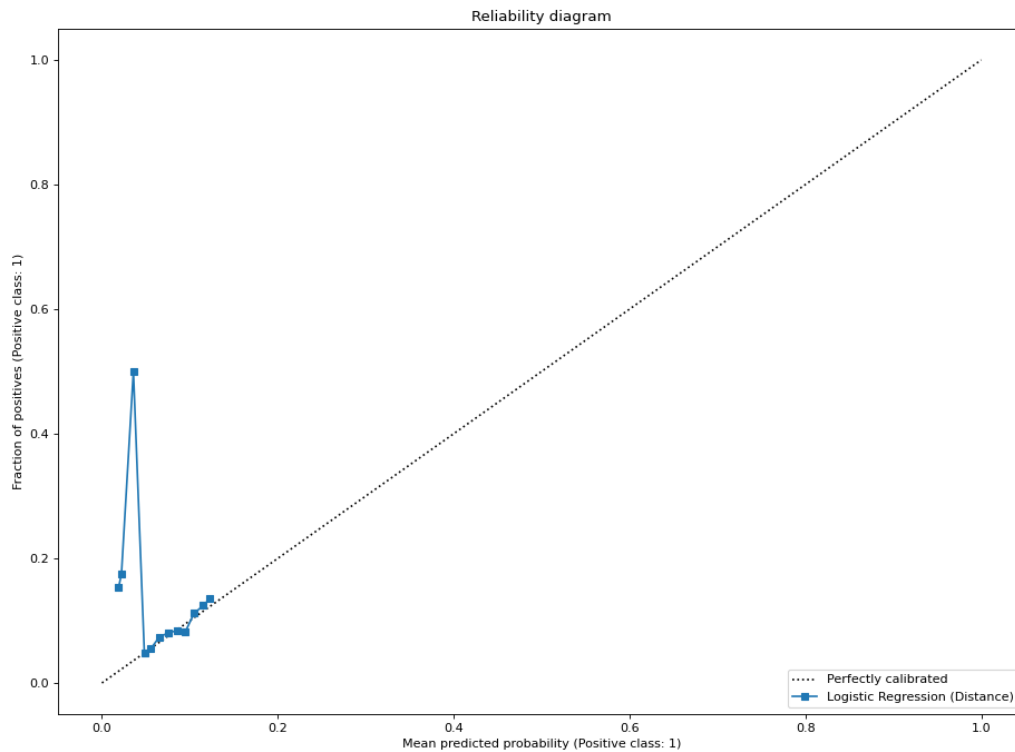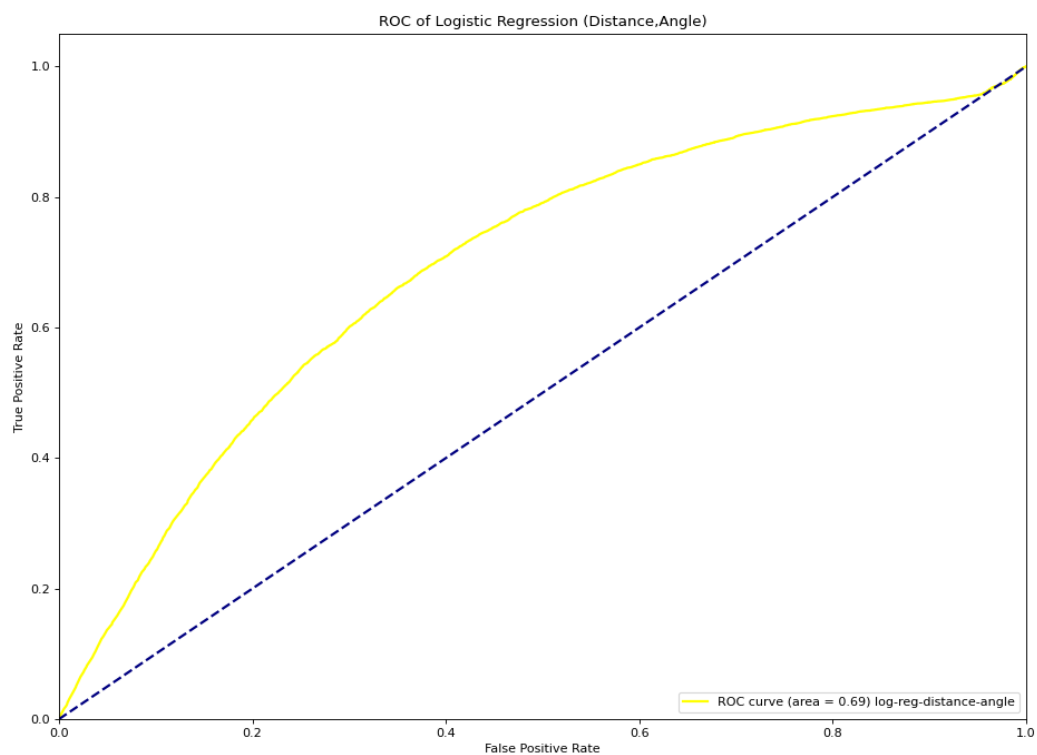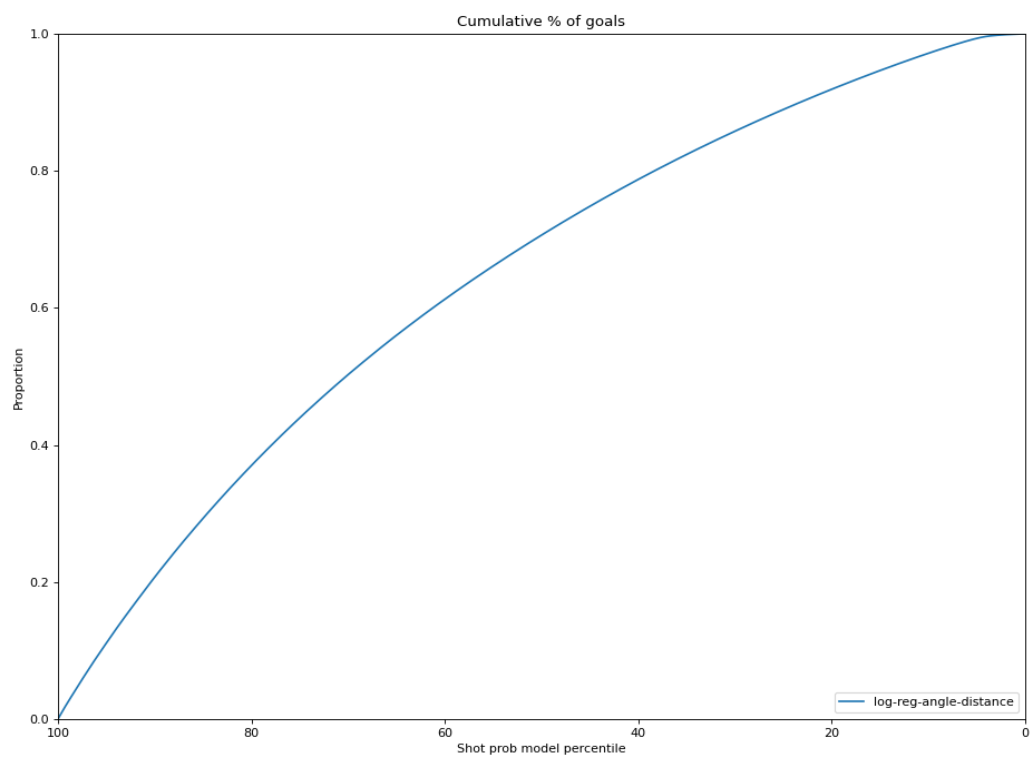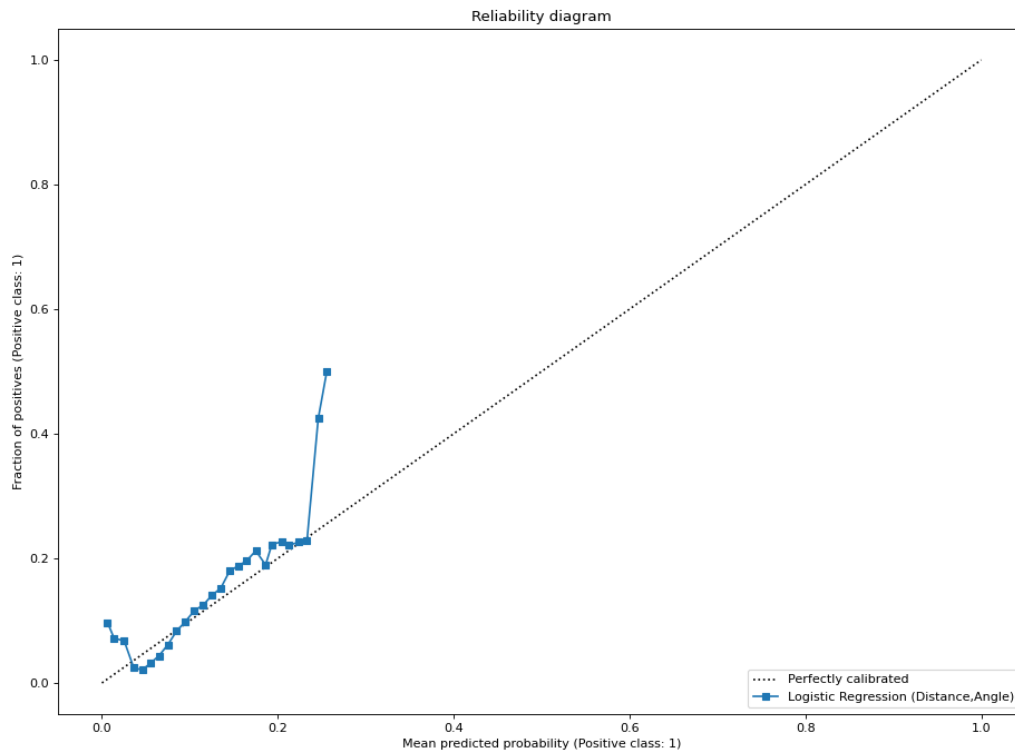
ROC of Logistic Regression (Distance,Angle)

Cumulative % of goals

Reliability diagram

We quickly come to the conclusion that the logistic regression model consistently predicts goals, regardless of the features used. It implies that a more sophisticated model must be taken into account for this binary classification, which will be investigated in Tasks 5 and 6. We even contrast the non-learning statistical method with the log-reg.