|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Chat bot performance 1 | Chat bot performance 2 | Chat bot performance 3 | ground truth | utterance |
| + TP | + TP | + | + | Order one cheeseburger |
| + TP | + TP | + | + | One cheeseburger |
| - FP | + TP | + | + | Order one cheeseburger and one coke |
| + TP | + TP | + | + | Can I have one cheeseburger |
| + TP | + TP | + | + | I’d like one coke |
| - FP | + TP | + | + | Can I get a cheeseburger |
| - FP | + TP | + | + | I’ll have a cheeseburger |
| + TP | + TP | + | + | Delete one fish burger |
| - FP | - FP | - FP | + | Change cheeseburger to fish burger |
| - FP | - FP | + TP | + | I don’t want diet coke |
| - FP | + TP | + | + | That’s it |
| - FP | + TP | + | + | Cancel one cheeseburger |
| - FP | + TP | + TP | + | Cancel my order |
| + FN | - TN | - | - | Order for me |
| - TN | - TN | - | - | One bag |
| - TN | - TN | - | - | Where is the pick-up place? |
| + FN | + FN | + FN | - | Order later |
| + FN | + FN | + FN | - | One folk |
| * TN | - TN | - |  | I am hungry |
| * TN | - TN | - |  | Which is the best burger |
| + FN | + FN | + FN |  | I don’t want to order |
| + FN | + FN | + FN |  | I’d like food |
| - TN | - TN | - | - | When do you close |
| - TN | - TN | - | - | Can I pay by credit card |
| - TN | - TN | - | - | What is in the combo one |
| + FN | + FN | - TN |  | I’d like to leave |
| Recall=0.454 | Recall=0.687 | Recall=0.75 |  |  |
| Precision=0.384 | Precision=0.846 | Precision=0.923 |  |  |
| TPR=0.454 | TPR=0.687 | TPR=0.75 |  |  |
| FPR=0.533 | FPR=0.25 | FPR=0.111 |  |  |

1

0.8

0.6

True Positive Rate

0.4

0.2

0.8

0.6

1

0.4

0.2

0 

False Positive Rate

ROC CURVE

1

0.8

0.6

Precision Rate

0.4

0.2

0.8

0.6

1

0.4

0.2

0 

Recall Rate

PR CURVE