

# Machine Learning Task: *Call Subcase Prediction*

ObserveAI

## 1 Task Definition

Given ASR generated transcripts for a set of call recordings, predict the correct subcase for each call. A subcase is defined as a general class of reason for calling the customer contact center.

The following is a list of possible subcases, along with their descriptions:

1. **customer:** the call is made by the customer regarding a non-order related issue, such as a login issue or a discount offer confirmation
2. **order:** the call could be made by the customer/shopper only regarding an order related issue, such as the issue of an incorrectly delivered order
3. **shopper:** the call is made by the shopper regarding a non-order related issue
4. **applicant:** the call is made for an application related issue
5. **misc:** the call is made for a reason not covered by the above categories

In the above descriptions, a *customer* is defined as a person placing a grocery order with the company. A *shopper* is a person completing the order on behalf of the company and for the customer. An *order* is a list of items required by the customer. And, an *applicant* is defined as a person applying to become a shopper for the company.

## 2 Data

### 2.1 Format

CSV formatted files with the following fields:

1. **call\_id:** unique id is given to each call
2. **text:** ASR output for the actual call. Contains a lot of noise, which is also a challenge in this task

3. `label`: The gold label subcase annotation for each call (missing from the evaluation dataset)

## 2.2 Datasets

1. `data_train.csv`: contains a total of 3460 call transcripts and can be used for training ML models.
2. `data_dev.csv`: contains a total of 1155 call transcripts and can be used for fine tuning the trained ML models and approximating the evaluation accuracy.
3. `data_eval.csv`: contains a total of 1157 call transcripts without the ‘label’ field value. The candidate’s applied ML skills will be evaluated on this dataset.

## 3 Evaluation

The evaluation metric to be used is average f1-score over the 5 classes defined in Section 1.

$$\text{average } f1\text{-score} = \frac{1}{n} \sum_{i=1}^n f1\text{-score}_i$$

where,  $n$  is the number of classes in dataset.

An exact definition of f1-score can be checked here: [https://en.wikipedia.org/wiki/F1\\_score](https://en.wikipedia.org/wiki/F1_score).

## 4 Submission

The candidate can assume this task to be complete upon the submission of the following:

1. The complete source code and instructions to run the modules.
2. A CSV file containing subcase prediction results for the set of transcripts present in the evaluation dataset (`data_eval.csv`). The file should be in the following format:

```
call_id, label
8721845, customer
8648189, customer
8066801, shopper
...
```

**Note:** The candidate is free to use any programming language for this task. The candidate is also allowed to use any open-source library for sub-tasks such as data pre-processing, and training and evaluating ML algorithms.

## 5 Final Note

In case of any clarification regarding the above task, feel free to drop a mail at `prabal@observe.ai` with `[ML Task]` as a prefix to the subject.