



C

Inference

Total inferred proteins: 918  
Rows with Label '1' > 0.75: 423 (46.08%)  
Rows with Label '1' between 0.25 and 0.75: 173 (18.85%)  
Rows with Label '1' < 0.25: 322 (35.08%)  
inference\_TT\_TFBS-L1O\_epochs\_120\_lr\_0.0000100\_model\_RBP\_TSTL\_small\_batch\_25\_exclude\_1.csv  
Total inferred proteins: 918  
Rows with Label '1' > 0.75: 429 (46.73%)  
Rows with Label '1' between 0.25 and 0.75: 90 (9.80%)  
Rows with Label '1' < 0.25: 399 (43.46%)

Inference completed...

PRONTO-TK Configuration

### UNIPROT

**go\_ids**  
FORMAT: python LIST of strings ["GO:<goid>"] - GO terms to be searched. You will get results of proteins matchin ANY of these GO Terms  
["GO:0003700"]

**go\_includeddescendants**  
FORMAT: string "True"/"False" - a flag to control whether the descendants of the GO terms have to be included in the search  
"True"

**go\_batchsize**  
FORMAT: -1 or <integer> - If -1, the "stream" REST service is used, otherwise the "search" REST call is used. "search" can be used to download in batches when the connection with UniProt is problematic or slow. In that case, 500 is a good value  
500

**go\_maxproteinsdownload**  
FORMAT: -1 means NO LIMIT, otherwise the MAX number of TOTAL proteins to be downloaded - The minimum is in any case batchsize.  
15000

**go\_taxonomies**  
FORMAT: python LIST of strings ["tax\_1", "tax\_2", ...]: Taxonomy IDs to be used for the search  
["1783272"]

**createflag**  
FORMAT: "yes"/"no"/"ask" - a flag to control whether a new dataset (or new embeddings) have to be generated if already present  
"ask"

**go\_folder**  
FORMAT: <string> - The folder where the files downloaded from Uniprot are stored and the embeddings will be created  
"Original Input//TFDNABind"

**datasetname**  
FORMAT: <string> - The name of the dataset to be created  
"tfbs"

### UNIPROT-Label 1

Save Save As