最終更新日：2025年2月05日

【機械学習実行手順】

1. 入力ファイル作成

作成手順は以下の通り

1. 機械学習データ取り出し.txt内のクエリを実行する．問題単位であれば「機械学習例題WID8」，単語単位であれば「機械学習単語単位」を実行．実験ごとのWIDとUIDは2019静大実験～～とかに書いている．
2. 結果が出たら画面一番下の「クエリ操作結果」→「エクスポート」を選択
3. 「エクスポート方法：簡易」，「フォーマット：CSV for MS Excel」を選択，それ以外はそのままにして「実行」を押す
4. ファイルがダウンロードされるので，1行に1種類のデータが入るようにする．

（落としてきた状態だと文字列が2行くらいに入っているだけなので，Excelの「データ」タブ→「区切り位置」を利用．区切り文字をセミコロンに指定し実行）

1. 問題単位の場合のみ　ファイルを2つに分割する．

分割の際はあるユーザ・ある問題のデータが2つのファイルに分かれないように，問題が変わるところで切る．上記をふまえて大体2等分くらいにすればOK．

1. ファイルの名前をinput～に変更して完成

※2のファイルが実行できないときは大体このファイルの作成がうまくいってない時．その時は大体③～⑤あたりでミスってるので，先輩が作ったデータと見比べてどこがおかしいか探したらまあ解決する．

※機械学習ではなくデータ分析の為にadminからデータをエクスポートする際は，③で「エクスポート方法：詳細」を選択後「フォーマット特有のオプション：1行目にカラム名を追加する」にしとくとデータが見やすい．

1. generateParametersFor〜.py を実行

機械学習のパラメータを生成するためのファイル

問題単位の時はForQuestion、単語単位の時はForWordを使う．

ForQuestion\_bはUターン回数に対して閾値（border）が設定されたもの（ForQuestionは閾値なし）．

現状は閾値5pxを採用（実験したらこの値が1番目視での計測回数に近かった）．

1. classify\_crossval\_fromCSV(〜.pyを実行

生成されたパラメータを用いて機械学習を実行するファイル

問題単位の時は(Question)，単語単位の時は(Word)を利用．

追加分

sampleSHAP:実際のLMSで使用している機械学習のプログラム

generata~ForWord\_update:迷い有りと無しがこれまではTRUEとFALSEと出力されていたのを問題単位と同様に2,3,4で出力されるように変更

（2:迷い有り，3:少し迷った，4:ほとんど迷わなかった）

研究室データベース: <http://153.120.0.32/phpmyadmin/index.php>