PPFST Crowd Methodology

The PPFST Crowd approach is to take a “Wisdom of Crowds” approach and use human beings to consider the merits of other selected models and vote for the “best model” – in terms of what is most likely to yield the most “points”. Their selection is based on a number of factors including information provided on the current and previous influenza seasons and the performance of models in the current and previous seasons. Crowd members are welcome to use any other sources they desire in their selection of the “best model”. The PPFST Crowd model provides probabilities for the bins associated with the seven targets of the Influenza Challenge at the National and Regional level.

This team is most interested in the predictions of the targets at the National level.

Volunteers are sought from the Pandemic Prediction and Forecasting Science and Technology (PPFST) Working Group. However, the response from this group was limited so additional crowd members were added from those who had previously or currently participating in the CDC challenge, colleagues of PPFST members, and family members.

Each week, a worksheet is provided to crowd members with information on the influenza season and the performance of models to select from. An epidemiological curve of weighted ILI% from all seasons (except 2009-2010) and the current season, based on <https://gis.cdc.gov/grasp/fluview/fluportaldashboard.html>, is provided at the National level and each region where votes are being collected. The worksheet provides the ILI% of the most recent weeks and asks the crowd member what they think the value/range will be for each target/location.

A ballot is provided on Tuesday, after submitted models are made available on Tuesday mornings from <https://github.com/cdcepi/FluSight-forecasts> and are returned prior to 9:00 pm on Tuesday night. Votes are compiled and submitted prior to midnight on Tuesday. The ballot shows the probability distribution of multiple models for each target. For each target on weighted ILI%, the ballot indicates that “the forecast score is related to the probabilities +/- 0.5% of the bin which contains the actual value”. For each target based on identifying the week, the ballot indicates that “the forecast score is related to the probabilities +/- 1 week of the bin which contains the actual value”. For all targets, the ballot asks “Which model do you think will get the highest score?” Initially, the model will be provided in the form of a Word document. It is hoped that the ballot will be available online, <https://jeffmorgan.shinyapps.io/>, at some point during the challenge.

Because there are 7 targets for 11 locations, there are 77 overall targets to consider. To reduce fatigue of crowd members, several simplifications are undertaken. Crowd members are asked to select targets for the US National level and no more than two other regions, typically Region 3 and Region 4. Most participants live in Region 3 or Region 4 and may understand the influenza behavior in that Region better than in more distant regions. Furthermore, the ballot allows for participants to use their selections for the National Level to apply to Regional forecasts or for their model selections for the first region apply to the second reason. For all members, the weights for the remaining eight regions are an average of the regions for the two regions where they voted on a model.

For the National location and regions for which there were votes, the weight *wi(Mj, Tk, Ll)* for each location *L*, target *T*, and model *M* is simply the number of votes for that model, target, and location divided by the total number of voting crowd members. For regions where votes were not collected, the weight is simply the average of the weights of the regions where votes were collected. Let *C* be the number of Crowd members who voted, *Rv* be the number of regions where votes were cast, and *M* be the number of Models to choose from.

For locations where votes are collected,

For locations where votes are not collected,

The probability assigned to each bin b for each target T and location L is simply the weighted average of the probabilities *pb(Mj, Tk, Ll)* for the bins of component models.

To further save time and avoid fatigue, crowd members may simply submit the worksheet with a value or range for each target and an algorithm will select the optimal model based on which model will yield the highest score.

When only one crowd member is able to respond (e.g., Thanksgiving holiday), that crowd member will provide two ballots.

Crowd members will be provided with an update on the scores of their predictions.