

Прогнозирование макроэкономических и финансовых переменных

В рамках данного хакатона участникам предстоит разработать модель, которая сможет быстро адаптироваться к задаче прогнозирования экономических переменных квартальной и месячной периодичности, которые модель не видела до этого.

Обучающая выборка

Обучающая выборка состоит из 69 экономических переменных месячной и 39 переменных квартальной периодичности в период с января 2003 года по декабрь 2015 года¹ и содержится в файле Train.xlsx. Месячные данные расположены на листе «Monthly», квартальные на листе «Quarterly». Даты расположены в первом столбце и состоят из года и месяца (например, 2003m01 – январь 2003 года). При этом квартал обозначается последним месяцем, т.е. I квартал 2003 года имеет обозначение 2003m03. Столбец для каждого ряда состоит из его названия, поля с преобразованием и значений самого ряда. Строка с преобразованием является справочной и содержит либо пустое значение, если ряд не преобразовался, либо «ln», если от исходного ряда был взят логарифм.

Тестовая выборка

Тестовая выборка состоит из 4446 файлов в каждом из которых отдельная задача прогнозирования не более чем 10ти временных рядов на не более чем 15 месяцев/5 кварталов (длина рядов не менее 5ти и не более 20ти лет). Файл может содержать только квартальные, только месячные, либо же показатели и с месячной и с квартальной периодичностью. Все переменные в тестовых файлах обезличены и имеют на каждом из листов названия «Var1», «Var2» и т.д. Все файлы форматированы таким образом, что месячные данные (при наличии) начинаются с даты 01m01, а квартальные с 01m03 (при наличии). При этом это абсолютно не означает, что данные начинаются с 2001 года. Год и месяц/квартал старта датасета может быть любым. Ячейки, которые необходимо спрогнозировать содержать надпись «Forecast».

Тестовые файлы содержат задачи по структуре похожи на те, которые возникают у экономистов при прогнозировании в реальности. Три основных типа прогноза, которые используются:

1. *Безусловный прогноз.* Стандартный прогноз, который подразумевает, что все переменные имеют одинаковую длину и прогноз осуществляется на одинаковое количество периодов вперед. В файле Test_example1.xlsx приведен пример, где необходимо спрогнозировать индекс потребительских цен (Var1 на листе «Monthly»), процентную ставку межбанковского рынка MIACR (Var2 на листе «Monthly»), цену на

¹ Некоторые переменные имеют первое значение позднее января 2003 года.

нефть марки Urals (Var3 на листе «Monthly»), реальный оборот розничной торговли (Var4 на листе «Monthly») и реальный ВВП (Var1 на листе «Quarterly»).

2. *Условный/сценарный прогноз.* В реальности экономистов зачастую интересует вопрос, что произойдет с экономикой в том или ином сценарии, например, если цены на нефть будут высокими или же наоборот низкими. Для таких ситуаций тестовые примеры не требуют прогноза для переменных, по которым задаются сценарии. Пример с такой задачей, где в качестве сценарной переменной используется цена на нефть представлен в файле Test_example2.xlsx (остальные переменные такие же как в предыдущем примере).
3. *Прогноз с задержкой выхода официальной статистики.* Расчет многих статистических показателей, таких как ВВП или оборот розничной торговли, требует некоторого времени, поэтому в момент прогноза часть переменных может быть недоступны за последние месяцы или кварталы. В таких случаях возникают задачи прогнозирования, где длина временных рядов одинаковой частотности может различаться. Примеры таких задач для безусловного и условного прогноза приведены в файлах Test_example3.xlsx и Test_example4.xlsx, где переменные реального оборота розничной торговли и реального ВВП выходят с задержкой.

В дополнительных файлах (Test_exampleXX.xlsx) также приведены примеры для проверки того, что модель учитывает различные типы тестовых задач и может с ними работать.

Выходные файлы

Входные файлы имеют название Test_input_XX.xlsx, где XX принимает значения от 1 до 4446. В качестве ответа к каждой задаче ожидается получить файл с названием Test_output_XX.xlsx, где будут заполнены все ячейки со значением «Forecast».