

Общий алгоритм вычисления агрегированного прогноза временных рядов

Д.С. Яшин¹, В.С. Мошкин¹, И.А. Мошкина¹, Н.Г. Ярушкина¹

¹Ульяновский государственный технический университет, Северный Венец 32, Ульяновск, Россия, 432027

Аннотация

В работе представлен общий алгоритм вычисления агрегированного прогноза временных рядов (ВР), в рамках которого применяются методы машинного обучения для настройки параметров гибридной комбинированной модели прогнозирования. Также приведены результаты экспериментов по применению разработанного алгоритма с использованием ВР соревнований «Computational Intelligence in Forecasting» (CIF).

Ключевые слова

Временные ряды, прогнозирование, агрегированный прогноз, машинное обучение

1. Введение

При проектировании технических систем (ТС) их состояние часто описывается временными рядами (ВР). Для принятия проектного решения необходимо прогнозировать будущее состояние объекта проектирования. В разработанном подходе применяются методы машинного обучения для настройки параметров гибридной комбинированной модели прогнозирования ВР. Согласно применяемому в данной работе предикативному подходу, необходимо определить набор значимых для рассматриваемой задачи характеристик ВР, совокупность которых достаточно полно описывает его динамику [1]. На рисунке 1 представлено схематическое описание алгоритма вычисления агрегированного прогноза ВР.

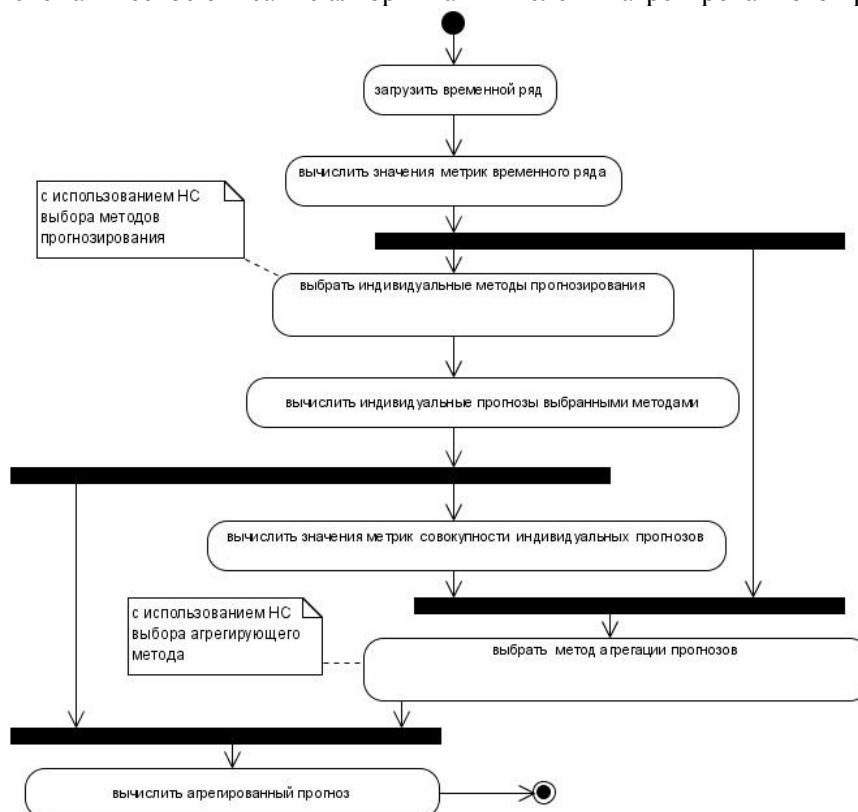


Рисунок 1: Алгоритм вычисления агрегированного прогноза временных рядов

Для проверки эффективности разработанного решения выбраны временные ряды с соревнований «Computational Intelligence in Forecasting» (CIF) [2].

- Первая бенчмарка CIF содержит 91 временной ряд различной длины (от 12 до 1089 наблюдений) и различной частоты наблюдений: день, месяц, квартал, год.
- Вторая бенчмарка CIF включает 72 временных ряда с частотой – месяц и длиной от 28 до 120 наблюдений.

Было проведено пять идентичных по алгоритму экспериментов, после чего по ним был получен усредненный итоговый результат. В ходе каждого эксперимента набор из 152 временных рядов бенчмарки разделялся случайным образом на обучающую (142 временных ряда) и контрольную части (10 временных рядов).

Из общей таблицы обучающей выборки исключались ВР контрольной части, после чего по оставшимся данным обучалась нейронная сеть выбора методов прогнозирования (с автоматическим подбором оптимального числа нейронов). Затем полученная нейронная сеть использовалась для выбора 8 лучших методов прогнозирования из базового набора для временных рядов контрольной части общей таблицы [3].

На рисунке 2 представлена диаграмма, отображающая по каждому из пяти экспериментов среднюю ошибку SMAPE по лучшим методам и SMAPE по всем методам.

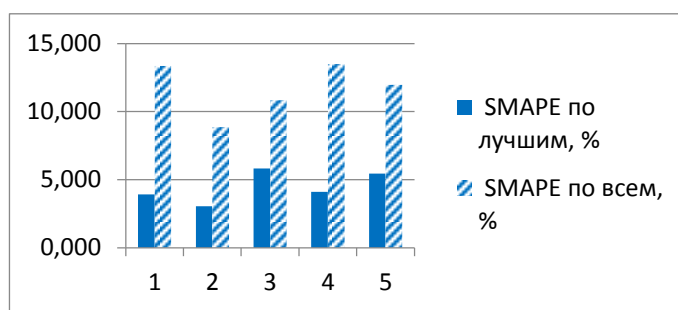


Рисунок 2: Результаты проведенных экспериментов

2. Заключение

Как видно из результатов экспериментов, применение нейронной сети выбора методов прогнозирования позволило в среднем по всем экспериментам снизить ошибку на 7,225%. Средняя ошибка по восьми выбранным нейронной сетью методам прогнозирования оказалась ниже средней ошибки по всем методам в 47 случаях из 50 (94%).

3. Благодарности

Работа выполнена при финансовой поддержке РФФИ, гранты № 19-07-00999 и 20-07-00672.

4. Литература

- [1] Yashin, D. An Approach to the Selection of a Time Series Analysis Method Using a Neural Network / D. Yashin, I. Moshkina, V. Moshkin // Computational Science and Its Applications – ICCSA. Lecture Notes in Computer Science. – 2020. – Vol. 12249. – P. 692-703. DOI: 10.1007/978-3-030-58799-4_50.
- [2] Computational Intelligence in Forecasting (CIF) [Электронный ресурс]. – Режим доступа: <http://irafm.osu.cz/cif/main.php>.
- [3] Romanov, A. Discrete and fuzzy models of time series in the tasks of forecasting and diagnostics / A. Romanov, V. Voronina, G. Guskov, I. Moshkina, N. Yarushkina // Axioms. – 2020. – Vol. 9(2). – P. 49.