

МАШИННОЕ ОБУЧЕНИЕ В ЗАДАЧЕ ПРОГНОЗИРОВАНИЯ ВРЕМЕННЫХ РЯДОВ МПЧ КОРОТКОВОЛНОВЫХ КАНАЛОВ РАДИОСВЯЗИ

Конкин Никита Александрович

*ФГБОУ ВО Поволжский государственный технологический университет,
студент аспирантуры 2-го курса, гр. ФА-21,
направление Физика и астрономия*

Пасова Анастасия Дмитриевна

*ФГБОУ ВО Поволжский государственный технологический университет,
студент магистратуры 2-го курса, гр. ИТСм-21,
направление Инфокоммуникационные технологии и системы связи*

Аннотация. Статья посвящена созданию алгоритма долгосрочного прогнозирования значений МПЧ. В работе проанализировано влияние различных методов обработки сырых значений максимально применимой частоты на результаты выполнения алгоритмов машинного обучения, таких как линейная регрессия и XGBoost. В качестве приемов обработки применялась фильтрация методом Савицкого — Голея и алгоритм изолированного леса для определения выбросов для суточного хода МПЧ.

Annotation. The article is devoted to the creation of an algorithm for long-term prediction of the values of the MPR. The paper analyzes the influence of various methods of processing raw values of the maximum applicable frequency on the results of machine learning algorithms, such as linear regression and XGBoost. As processing techniques, the Savitsky — Goley filtration method and the isolated forest algorithm were used to determine emissions for the daily course of the MPR.

Ключевые слова: радиоканал, долгосрочное прогнозирование, МПЧ, машинное обучение, линейная регрессия, XGBoost, изолированный лес.

Keywords: radio channel, long-term forecasting, MHR, machine learning, linear regression, XGBoost, isolated forest.

Введение. Более эффективно использовать ресурсы радиопередающей системы позволяет определение частотных границ коротковолнового канала связи [1,2]. Граничные значения частот изменяются из-за неустойчивого во времени состояния ионосферы, поэтому оптимальную рабочую частоту выбирают на 15% ниже максимально применимой частоты (МПЧ) [1]. Целью работы является создание алгоритма долгосрочного прогнозирования (от нескольких часов до нескольких суток) значений МПЧ с использованием методов машинного обучения. Данный алгоритм реализован на языке программирования Python, для сравнения используется метод линейной регрессии и XGBoost [3]. Выбор линейной регрессии обусловлен простотой самой модели и низким потреблением вычислительных ресурсов. Модель XGBoost взята как более эффективная, но потребляющая большее количество ресурсов. Обучение алгоритма происходит на временном ряду, который формируется с помощью функции TimeSeriesSplit библиотеки sklearn, а аппроксимация временных значений реализована за счет полиномиального фильтра Савицкого — Голея [4].

Задача исследования состояла в оптимизации метода фильтрации временных рядов путем определения окна фильтрации и порядка фильтра, выборе оптимального способа деления всего временного ряда МПЧ на учебные и тестовые выборки. Критерием качества настройки алгоритма являлись коэффициент детерминации и среднеквадратичное отклонение используемых методов машинного обучения от экспериментальных значений МПЧ, полученных в результате диагностики ЛЧМ ионозондом ВЧ радиоканалов связи [1,2].

При проведении эксперимента была проанализирована автокорреляционная функция (АКФ) для каждого суточного хода МПЧ, в ходе чего была установлена длина окна фильтрации как 95%-й доверительный интервал, в который попадают отсчеты АКФ. Порядок фильтра определялся из условия 95% взаимно корреляционной функции ряда МПЧ до и после фильтрации.

Результаты эксперимента представлены на рисунках 1 и 2. На графиках изображены усредненные результаты за пять выборок для каждого соотношения: размер окна фильтрации/порядок фильтра.

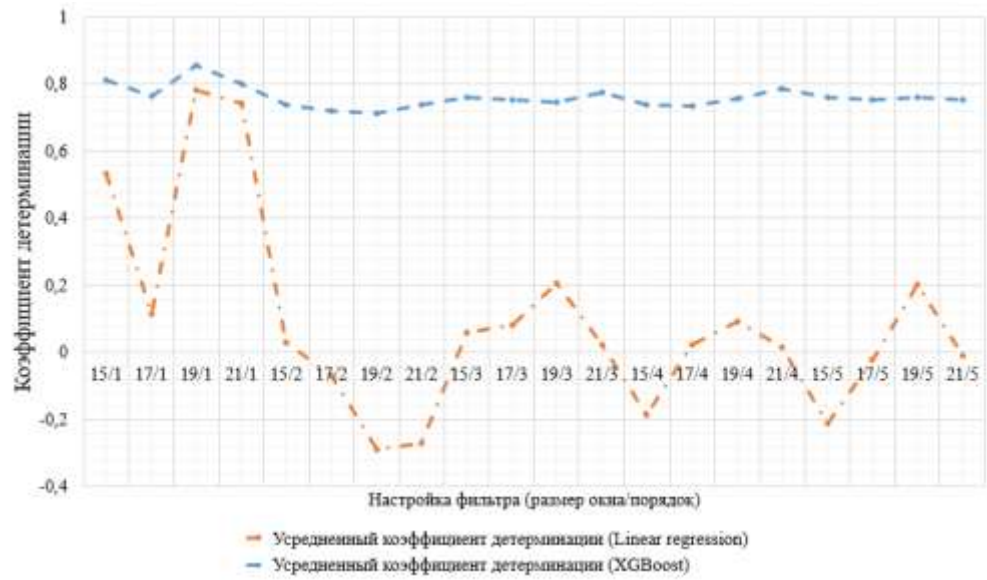


Рисунок 1. Зависимость коэффициента детерминации от показателя окно/порядок.

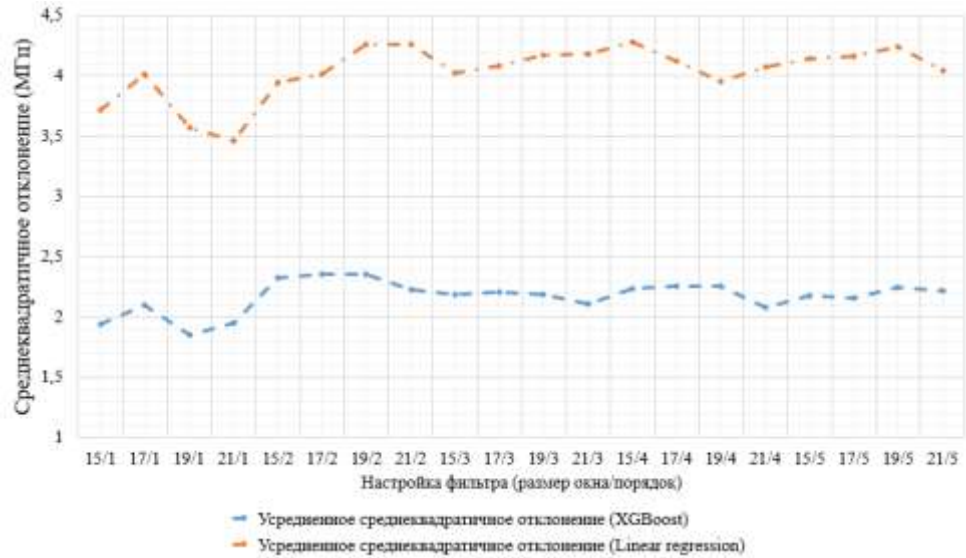


Рисунок 2. Зависимость СКО от показателя окно/порядок

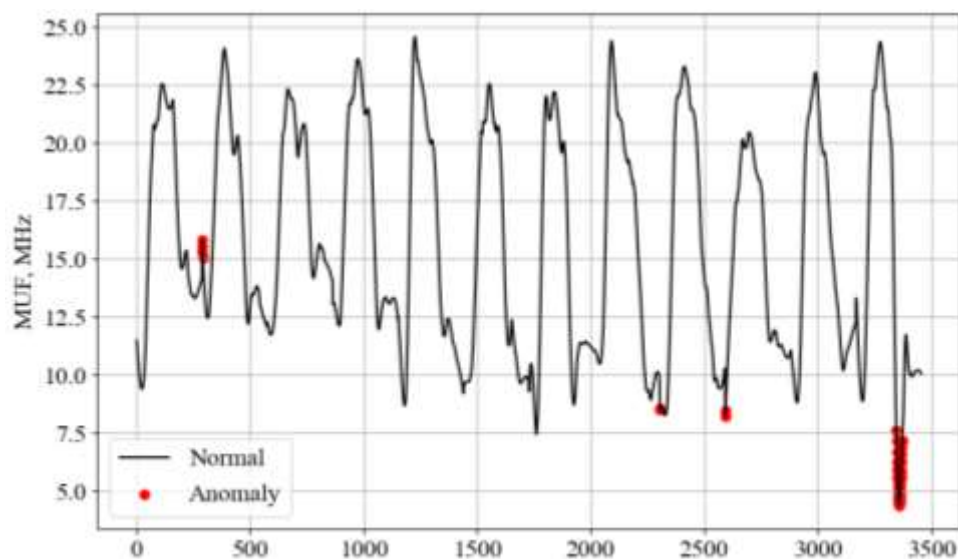


Рисунок 3. Суточный ход МПЧ после фильтрации за 12-ть дней. Точками отмечены протектированные выбросы.

Для более корректной работы алгоритмов машинного обучения важны определение и фильтрация выбросов временной выборки. В работе для решения данной задачи выбран алгоритм изолированного леса библиотеки sklearn. Данный алгоритм возвращает коэффициент «аномальности» для каждого временного отсчета, затем исследователем выбирается пороговое значение этого коэффициента, которое подбирается в зависимости от уровня выбросов во временном ряду. На рисунке 3 линией показана зависимость суточного хода МПЧ для 12-ти дней, а точками отмечены отсчеты ряда, которые относятся к разряду выбросов. Полученные «аномальные» впоследствии подвергаются сглаживанию.

Результаты работы алгоритмов машинного обучения представлены на рисунке 4, для случая, когда среднеквадратичное отклонение модели XGBoost равно 1,8 МГц, линейной регрессии 3,2 МГц, а коэффициенты детерминации равны 0,83 и 0,37 соответственно.



Рисунок 4. Результаты выполнения алгоритмов машинного обучения по прогнозированию суточного хода МПЧ (примерно 1.3 суток).

Заключение. На основе полученных данных определено оптимальное соотношение окно–порядок, равное 19/1, по двухкритериальному показателю: максимума коэффициента детерминации 0,86 и наименьшему значению СКО в 1,8 МГц для модели XGBoost. Результаты применения линейной регрессии в данной точке также являются оптимальными, т.е. имеют максимальное значение коэффициента детерминации и минимальное СКО, но проигрыш модели XGBoost в точности составляет примерно 1,5 МГц. Применение алгоритма изолированного леса для обработки выбросов позволило повысить точность моделей в среднем на 1,5 МГц по

СКО и 0,2 пункта по коэффициенту детерминации. Таким образом, для дальнейших исследований метода долгосрочного прогнозирования МПЧ наиболее эффективным является подход модели XGBoost.

СПИСОК ЛИТЕРАТУРЫ:

1. Иванов В.А., Иванов Д.В., Рябова Н.В., Лыонг В.Л., Рябова М.И. Синтез, анализ и прогнозирование характеристик ионосферных линий декаметровой радиосвязи. Минобрнауки РФ. "МарГТУ". Йошкар-Ола, 2011.
2. Ivanov V., Ivanov D., Ryabova N., Ryabova M., Chernov A., Elsukov A. Algorithms and software of whitening of noise in the reception of broadband chirp signals in the HF communication channel // Journal of Applied Engineering Science. 2015. Т. 13. № 4. С. 245-250. DOI:10.5937/jaes13-9342
3. XGBoost Tutorials URL: <https://xgboost.readthedocs.io/en/latest/tutorials/index.html> (дата обращения: 04.10.2021).
4. savgol_filter // scipy URL: https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal.savgol_filter.html (дата обращения: 01.10.2021).