

Модель градиентного бустинга для оценки агрометеорологических условий вегетационного периода озимой пшеницы.

Научный руководитель – Асауляк Ирина Федоровна

Салмин Андрей Сергеевич

Аспирант

Российский государственный аграрный университет МСХА имени К.А. Тимирязева,
Агрономии и биотехнологии, Земледелия и агрометеорологии, Москва, Россия
E-mail: ex1.95@mail.ru

Методология количественной оценки агрометеорологических условий вегетационного периода сельскохозяйственных культур, как правило, базируется на степени отличия расчетной ожидаемой урожайности от некоторой эталонной величины [1]. В соответствии с этим подходом, урожайность представляет собой интегральную характеристику сложившихся агрометеорологических условий за определенный временной период. Для получения наиболее достоверной оценки требуется усовершенствование методов прогноза формирования урожайности на основании метеорологических данных.

Особенностью ансамблевых моделей является использование нескольких обучающихся алгоритмов с целью повышения лучшей эффективности прогнозирования, чем можно было бы получить от каждого обучающего алгоритма по отдельности. Технология построение модели градиентного бустинга базируется на подходе, который заключается в создании последовательных композиций моделей прогнозов, где каждая новая модель компенсирует недостатки полученных результатов на предыдущем шаге [2, 3]. В отечественной агрометеорологической практике не имеется достаточных исследований по применению технологий машинного обучения в решениях задач прогнозирования.

В настоящей исследовательской работе рассматривается подход к решению прогностической задачи с помощью градиентного бустинга над решающими деревьями. Модель градиентного бустинга для оценки агрометеорологических условий была реализована на языке Python с помощью прикладных научных библиотек XGBoost, Numpy. В качестве предикторов использовались индекс SPI, дефицит влажности воздуха, сумма активных температур полученные для территории Ростовской области. Оптимизатором модели выступал алгоритм Tree-Structured Parzen Estimators (ТРЕ). Для метрики оценки качества модели использовалась средняя абсолютная ошибка (MAE). На основании данного исследования была получена ошибка прогноза 4.3 ц/га.

Источники и литература

- 1) Сиротенко О.Д. Основы сельскохозяйственной метеорологии. Том II. Методы расчетов и прогнозов в агрометеорологии. Книга 1. Математические модели в агрометеорологии. // Учебное пособие. Обнинск, ФГБУ "ВНИИГМИ-МЦД", – 2012 –136 с.
- 2) Горшенин А.К. Мартынов О.П. Гибридные модели экстремального градиентного бустинга для восстановления пропущенных значений в данных об осадках // Информатика и ее применение, 2019. Т. 13. Вып. 3. С. 34-40.
- 3) Мартин О. Байесовский анализ на Python / пер. с англ. А.В. Снастина. – М.: ДМК Пресс, 2020. – 340 с.