

Agriculturing

Статьи:

1. [Статья \(конкретно выживаемость\)](#): предложена глубокая нейронная сеть для прогнозирования урожая (time-to-harvest) томатов в умных теплицах (учитываются параметры: температура, влажность, освещенность и характеристики почвы). Данные: табличные, но нет в открытом доступе (можно запросить). Комментарии: используются совсем простые модели (CoxPH), качество не идеальное (CI точно).
2. [Статья \(конкретно выживаемость\)](#): прогнозирование времени до урожая леса (до вырубки) по данным FIA США (используются отобранные 14 переменных). Модели: RSF, CoxPH. Метрики: AUC(t), IBS(t) (за первые 5 лет хорошие метрики, а дальше не очень хорошо).
3. [Статья](#): построение системы управления умной теплицей в реальном времени (контролируется работа кондиционера, насосов и освещения на основе датчиков). Решение простое. Данные: не найдены.
4. [Статья](#): прогноз урожайности за 1-4 недели (дискретная сетка) с помощью нескольких LSTM моделей. Данные табличные (нет в открытом доступе), признаки логичные: кумулятивное состояние объекта/среды за последние N дней. Нет цензурирования.
5. [Survey-статья](#): обзор NN в прогнозе урожайности. Много статистики, есть обзор популярных источников данных/направлений исследований/методов (ключевые слова: crop yield prediction).
6. [Survey-статья](#): обзор ML в точном земледелии. Анализ важных признаков (+датчики), источников данных, применение ML для прогноза урожайности и обнаружения болезней.

Наборы данных:

1. Набор [SICKLE](#) снимков рисовых полей (Landsat-8, Sentinel-1, Sentinel-2) - многоспектральные, тепловые и микроволновые датчики за 2018-2021 года. [Статья](#): есть обзоры других наборов, решались 4 задачи (сегментация культур, прогноз времени урожая, прогноз урожайности, поиск аномалий посевов) Код: [GitHub](#)
2. [Табличный набор на kaggle](#) (+ JN): собран за 1961-2016 года (10 культур, 168 стран). Достаточно общий: параметры среды за год (температура, дожливость, пестициды, общая урожайность). Нет времени до урожая, нет индивидуальности.

3. [Набор \(изображения со спутника\) + Статья](#): собран за 2020-2023 годы в Эфиопии. Основная задача: бинарная классификация наличия урожая (но пишут, что разметка более глубокая). Потенциально хотели добавить временные ряды (мониторинг за область). Код: [GitHub](#) + [доступ к данным](#)
4. [Табличный набор на kaggle](#): USFS Forest Inventory and Analysis (FIA) - набор состояния лесной площади (виды, размер и состояние деревьев; общий рост деревьев, смертности и "урожай"/вырубка).
5. [Набор \(изображения со спутника\) + Статья](#) по фенологии кукурузы в Китае за 1985–2020 годы. Содержит ежегодные данные о фенологических стадиях кукурузы с пространственным разрешением 30 м, включая даты посева и сбора урожая. В открытом доступе не найден, но можно написать.
6. [Набор \(изображения объектов\) + Статья](#): Raspberry PhenoSet, набор изображений малины для оценки урожайности. Содержит отдельные изображения малины на различных стадиях развития для задач прогнозирования урожая и времени сбора. Есть таблица существующих наборов изображений (с ссылками). Минусы: сложно экстраполировать прогноз отдельных "саженцев" на поле.

Можно посмотреть:

1. [IOT-Smart_Farming](#) - сбор данных с помощью датчиков (временные ряды). Возможно какие-то данные накоплены (снимаются параметры среды, свойства почвы)
2. Набор данных [IFPRI](#) и связанные [работы](#)
3. Набор данных [FAOSTAT](#)