Сравнительная оценка моделей глубокого обучения для классификации закодированных последовательностей сайтов сплайсинга

Источник: Frontiers in AI — Medicine

Оригинал: https://www.frontiersin.org/articles/10.3389/frai.2026.1886470

геномика | глубокое обучение | классификация | сплайсинг

Введение

Вычислительные классификаторы могут быть протестированы на публичных закодированных наборах данных по сайтам сплайсинга, однако оценка производительности будет зависеть от представления входных данных, конвейера предобработки, конфигурации модели и дизайна оценки.

Методы

В данном исследовании изучается производительность 15 различных архитектур — сверточных, гибридных и глубокого обучения — на публичном закодированном наборе данных по сайтам сплайсинга для задачи классификации на 3 класса. Результаты основных моделей были сохранены из единого рабочего процесса с разделением на обучающую, валидационную и отложенную тестовую выборки и представлены как иллюстративные эффекты, а не как статистически значимое доказательство превосходства.

Результаты

Наивысшая точность на отложенной тестовой выборке (96,87%) была получена с помощью пользовательской сверточной нейронной сети (Custom CNN) среди сохраненных конфигураций. MLP Mixer показал наивысшую

точность в процессе обучения (99,90%), однако модель показала более низкие результаты на валидационной выборке (93,33%) и отложенной тестовой выборке (93,73%), что указывает на некоторое переобучение на обучающих данных. В работе представлены сведения о емкости моделей, соотношении параметров к объему выборки, кривых обучения, метриках по классам, а также целевое сравнение взвешивания классов. Взвешивание классов выполнялось только на этапе обучения, однако для протестированной конфигурации Custom CNN взвешивание классов не обеспечило наивысшую производительность для наблюдаемого набора данных.

Обсуждение

Выводы ограничены результатами классификации, которые могли быть вычислены для закодированного эталонного набора данных и описанной процедуры оценки. Результаты не демонстрируют статистически значимого превосходства моделей в смысле внешней обобщаемости, клинической диагностической валидности, обнаружения мутаций на уровне пациентов, клинической применимости или готовности к внедрению. Для более общих выводов потребовались бы повторные оценки без утечек данных, внешние наборы данных, сохраненные предсказания моделей на выборках, статистическое сравнение выходных данных моделей и клинически валидированные данные на уровне пациентов.