УДК 004.052.2

РАЗРАБОТКА СЕРВИСА ДЛЯ ПОЛУАВТОМАТИЧЕСКОГО ТЕСТИРОВАНИЯ НЕЙРОСЕТЕВЫХ МОДЕЛЕЙ И ОЦЕНКИ ИХ РОБАСТНОСТИ В ОБЛАСТИ ЗДРАВООХРАНЕНИЯ

Змиевский Д.А. федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики»,
Полевая Т.А. федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики»
Научный руководитель – кандидат технических наук, старший научный сотрудник Гусарова Н.Ф.
федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики»

Введение.

В современных реалиях системы с использованием искусственного интеллекта играют важную роль в различных сферах. Применение ИИ в медицине позволяет автоматизировать процесс первичного диагностирования различных заболеваний, повысить эффективность и точность принимаемых решений.

Однако, для полного внедрения подобных систем в реальную клиническую практику, где необходимы тщательная оценка эффективности и объяснимость полученных результатов, требуется унификация проверки разработанных решений с целью установления способности моделей к обобщаемости и устойчивости к изменению модальности и характеристик данных из обучающего распределения.

Разработка сервиса для полуавтоматического тестирования решит проблемы в необходимости унификации оценки эффективности и робастности разрабатываемых нейросетевых моделей.

Основная часть.

Робастность нейросетевых моделей относится к их способности поддерживать высокую производительность и точность на различных входных данных, включая те, которые отличаются от данных, использованных для обучения модели. Оценку робастности разрабатываемых нейросетевых решений будет производится в три этапа: аугментация, Out-of-Distribution, Adversarial Data Augmentation.

Аугментация данных - это процесс создания новых образцов данных путем применения различных преобразований к существующим образцам, такие как повороты, масштабирование, сдвиги, отражения, изменение яркости и контрастности и другие [1]. Эти преобразования позволяют создавать разнообразные вариации исходных данных, что помогает моделям обучаться на более разнообразных и реалистичных примерах.

Out-of-Distribution (OOD) относятся к данным, которые значительно отличаются от тех, на которых была обучена модель машинного обучения. Другими словами, данные OOD поступают из распределения, которое находится за пределами диапазона распределения обучающих данных. Обнаружение и обработка таких данных важны, поскольку модели машинного обучения обычно разрабатываются для составления прогнозов в пределах

диапазона обучающих данных [2]. При работе с OOD данными эти модели могут выдавать ненадежные или неверные прогнозы.

Аугментация данных в контексте адверсиального обучения (Adversial Data Augmentation) относится к применению преобразований к обучающим данным с целью улучшения робастности моделей глубокого обучения против адверсариальных атак [3]. Адверсариальные атаки - это злонамеренные попытки внести незаметные изменения во входные данные с целью ввода модели в заблуждение и получения неправильных результатов.

Выводы.

Разработан прототип сервиса с методами тестирования робастности нейросетевых моделей.

Список использованных источников:

1. Alhassan Mumuni, Fuseini Mumuni, “Data augmentation: A comprehensive survey of modern approaches” // Array – Volume 16 – 2022 – 100258, ISSN 2590-0056.

2. J. Yang, K. Zhou, Y. Li, and Z. Liu., “Generalized Out-of-Distribution Detection: A Survey” // CoRR – abs/2110.11334 – 2021.

3. Mesut Ozdag, “Adversarial Attacks and Defenses Against Deep Neural Networks: A Survey” // Procedia Computer Science – Volume 140 – 2018 – Pages 152-161.

Змиевский Д.А. (автор) **Подпись**

Полевая Т.А. (автор) **Подпись**

Гусарова Н.Ф. (научный руководитель) **Подпись**