

Министерство образования Республики Беларусь
Учреждение образования
Белорусский государственный университет
информатики и радиоэлектроники

УДК 004.383

Ровдо
Николай Русланович

Фармакологическая система оценки терапевтической активности веществ

АВТОРЕФЕРАТ

на соискание академической степени
магистра информатики и вычислительной техники

по специальности 1-40 81 04 – Обработка больших объемов информации

Научный руководитель
Жвакина А.В.
к.т.н., доцент

Минск 2020

КРАТКОЕ ВВЕДЕНИЕ

С развитием информационных технологий и телекоммуникаций жизнь становится все более мобильной и информативной, новые технологии прочно входят в различные отрасли хозяйствования, сферы жизни и несут новые нормы в них. В связи с реформированием экономики, с взятием курса на её инновационное развитие, всё чаще и чаще в повседневной работе в большинстве отраслей начинают использовать различные средства информационно-вычислительной техники и соответственно программного обеспечения. Уровень развития медицинской отрасли в государстве – один из важнейших признаков ее технологического прогресса и цивилизованности.

Самые широко применяемые средства ИТ в наше время – это интернет, мобильные телефоны и компьютеры. Тем не менее, каждая узкая отрасль науки и производства имеет своё специально разработанное программное обеспечение, обеспечивающее работу систем и обрабатывающее информацию. Использование информационных технологий в медицине является логичным шагом, так как быстрый доступ к информации и обмен ею существенно сокращает временные издержки на поиск решений, а время часто является главным фактором в спасении жизни человека. Кроме того, применение технологий машинного обучения и нейросетей выводит её на качественно новую ступень.

Компьютер все больше используется в области здравоохранения, что бывает очень удобным, а порой просто необходимым. Благодаря этому медицина приобретает сегодня совершенно новые черты. Во многих медицинских исследованиях просто невозможно обойтись без компьютера и специального программного обеспечения к нему. Этот процесс сопровождается существенными изменениями в медицинской теории и практике, связанными с внесением корректив к подготовке медицинских работников. Современный период развития общества характеризуется сильным влиянием на него компьютерных технологий, которые проникают во все сферы человеческой деятельности, обеспечивают распространение информационных потоков в обществе, образуя глобальное информационное пространство. Они очень быстро превратились в жизненно важный стимул развития не только мировой экономики, но и других сфер человеческой деятельности. Трудно найти сферу, в которой сейчас не используются информационные технологии.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Цель и задачи исследования

Целью диссертационной работы является разработка программного обеспечения для обработки химических формул с целью решения задачи классификации данных по терапевтическому эффекту на человека.

Для достижения поставленной цели необходимо решить следующие задачи:

1. Определить возможности сверточных и классических нейронных сетей в задачах классификации в сфере машинного обучения, а также их эффективность.
2. Исследовать возможность и пользу предварительного преобразования входных данных с целью ускорения работы и улучшения эффективности системы.
3. Подобрать и предварительно преобразовать (при необходимости) обучающую выборку данных, предназначенную для настройки модели классифицирующей нейронной сети.
4. Разработать архитектуру нейронной сети, а также программной системы, способной обучиться распознавать терапевтический эффект на заданном наборе данных.
5. Реализовать ПО для распознавания терапевтического эффекта заданных химических формул с помощью обученной модели нейронной сети.
6. Провести экспериментальные исследования разработанной системы.

Объектом исследования являются нейронные сети глубокого обучения: сверточные и классические, а также машинные методы.

Предметом исследования является математическое и программное обеспечение компьютерных систем для решения задач многоклассовой классификации в машинном обучении.

Основной гипотезой, положенной в основу диссертационной работы, является возможность использования нейронных сетей для решения задач классификации химических структур на предмет наличия абстрактных признаков, в частности, принадлежности определённому терапевтическому эффекту. Предварительная обработка химических формул позволит повысить производительность классифицирующей модели, нейронные сети способны выявить множество абстрактных признаков у объекта. Полученная модель послужит средством анализа и поможет в скрининге при разработке лекарственных средств.

Личный вклад соискателя

Результаты, приведенные в диссертации, получены соискателем лично.

Апробация результатов диссертации

Основные положения диссертационной работы докладывались и обсуждались на 55-й юбилейной научной конференции аспирантов, магистрантов и студентов БГУИР (Минск, Беларусь, 2019);

Опубликованность результатов диссертации

По теме диссертации опубликована 1 печатная работа, находящаяся в сборниках трудов и материалов научных конференций.

Структура и объем диссертации

Диссертация состоит из введения, общей характеристики работы, трех глав, заключения, списка использованных источников и списка публикаций автора. В первой главе представлен анализ предметной области, выявлены основные существующие проблемы в рамках тематики исследования, показаны направления их решения. Вторая глава посвящена анализу программных средств, предназначенных для разработки нейронных сетей. В третьей главе предложена практическая реализация ПО для анализа и классификации, представлены результаты экспериментальных исследований и практического применения разработанной системы.

Общий объем работы составляет 102 страниц, из которых основного текста – 80 страниц, 44 рисунка на 29 страницах и список использованных источников из 40 наименований на 23 страницах.

ОСНОВНОЕ СОДЕРЖАНИЕ

Во **введении** определена область и указаны основные направления исследования, показана актуальность темы диссертационной работы, дана краткая характеристика исследуемых вопросов, обозначена практическая ценность работы.

В **первой главе** проведен анализ применяемых архитектурных решений систем, используемых при решении задач молекулярного докинга.

Вторая глава посвящена технологиям используемых при разработке программного продукта, среди которых Google Colab, Matlab, Tensorflow, Django, Ajax, Smiles, Word2vec, свёрточные нейронные сети и т.д.

В **третьей главе** рассмотрены задачи по исследованию проблемы оценки фармакологических групп по химическим формулам, а также реализация системы. Исследование состоит из первоначального обзора задачи и входящих данных, предложено несколько способов их интерпретации, для решения задач использовано и исследовано два подхода к решению задачи, проведён

сравнительный анализ и выдвинуто заключение по проделанной работе. В последней части показаны подробности реализации программного продукта, а также инструкция по его использованию.

ЗАКЛЮЧЕНИЕ

Основные научные результаты диссертации

В ходе работы над магистерской диссертацией была сформулирована и поставлена задача по разработке фармакологической системы оценки терапевтической активности веществ, которая позволит в большей степени оптимизировать процесс подбора комбинаторной библиотеки, а также упростить молекулярный докинг при разработке оригинального лекарственного средства.

Фармакологическая система была разработана с целью повышения производительности и качества процесса драг-дизайна, что позволит сделать его более эффективным.

Для того чтобы правильно спроектировать систему, необходимо было грамотно проработать предметную область, также разработать основной процесс предметной области. Это позволило подробно разобраться и понять, что должно выполнять приложение.

Важным этапом явилось проектирование такого интерфейса системы, чтобы подготовленному пользователю было привычно пользоваться приложением.

Система может быть улучшена за счет добавления алгоритмов докига. Такой функционал уместен при профессиональной и коммерческой разработке препаратов. Такой функционал может быть реализован в виде мобильного приложения, синхронизирующийся с основной шиной данных.

В ходе работы были проанализированы данные, а также набор методов обучения, для эффективной классификации. По итогам работы можно заявить, что структура химических связей в веществе имеет корреляцию терапевтического эффекта на человека.

СПИСОК ОПУБЛИКОВАННЫХ РАБОТ

1-А. Ровдо, Н.Р. Фармакологическая система оценки терапевтической активности веществ / Н.Р. Ровдо // Компьютерные системы и сети: материалы 55-ой научной конференции аспирантов, магистрантов и студентов. – Минск: БГУИР, 2019.

Библиотека БГУИР