Министерство образования Московской области

Государственное бюджетное профессиональноеобразовательное учреждение

Московской области «Физико-технический колледж»

ЗАЩИЩЕНО

ОЦЕНКА \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Руководитель специальности 09.02.07 Информационные

системы и программирование

Модель автоматической проверки подписи на основе сиамских нейросетей

Пояснительная записка к курсовому проекту

**по МДК 13.02 Технологии машинного обучения и анализа данных**

Руководитель КП Разработал

преподаватель студент группы ИСП 4-5

ГБПОУ МО «Физтех-колледж»

\_\_\_\_\_\_\_\_\_\_ Г. В. Базяк \_\_\_\_\_\_\_\_ А. Лим

**Министерство образования Московской области Государственное бюджетное профессиональное образовательное учреждение Московской области «Физико-технический колледж»**

|  |
| --- |
| Специальность 09.02.07  Информационные системы и Программирование Квалификация «Программист» |

**УТВЕРЖДЕНО**

цикловой комиссией специальности 09.02.07

«\_\_\_\_» \_\_\_\_\_\_\_\_\_\_\_\_\_\_ 2024 года

Руководитель специальности

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

ЗАДАНИЕ

Для курсового проектирования по МДК 13.02 «**Технологии машинного обучения и анализа данных**» студенту 4 курса группы 4-5

Лим Аллан

(фамилия, имя, отчество)

Тема задания: Модель автоматической проверки подписи на основе сиамских нейросетей.

При выполнении курсового проекта на указанную тему должны быть представлены

1. Пояснительная записка

Введение

1 Анализ предметной области и описание технологии проектирования ИC;

* 1. Описание предметной области и назначение системы
  2. Определение целевых групп пользователей
  3. Определение ограничений проектного решения

1.4 Анализ предметной области на уровне построения диаграммы использования системы (USE-CASE) и ER-диаграммы

1.4.1 Диаграмма вариантов использования (Use-Case)

1.4.2 ER-диаграмма (Диаграмма сущность-связь)

1.4.3 Вывод

2 Требования к системе

2.1 Требований к функциональности приложения

2.2 Требования к графическому интерфейсу

2.3 Требования к надёжности системы

Дата выдачи: « » 2024г. Срок окончания: « » 2024 г

Руководитель курсового проектирования Г.В. Базяк

СОДЕРЖАНИЕ

[Введение 4](#_Toc186044894)

[1 Анализ предметной области и описание технологии проектирования ИС 6](#_Toc186044895)

[1. 1 Описание предметной области и назначение системы 6](#_Toc186044896)

[1.2 Целевые группы пользователей 7](#_Toc186044897)

[1.3 Ограничения проектного решения 10](#_Toc186044898)

[1.4 Анализ предметной области на уровне построения диаграммы использования системы (USE-CASE) и ER-диаграммы 12](#_Toc186044899)

[1.4.1 Диаграмма вариантов использования (Use-Case) 12](#_Toc186044900)

[1.4.2 ER-диаграмма (Диаграмма сущность-связь) 13](#_Toc186044901)

[1.4.3 Вывод 14](#_Toc186044902)

[2 Требования к системе 16](#_Toc186044903)

[2.1 Требований к функциональности приложения 16](#_Toc186044904)

[2.2 Требования к графическому интерфейсу 16](#_Toc186044905)

[2.3 Требования к надёжности системы 16](#_Toc186044906)

# Введение

Документы играют ключевую роль в функционировании современного общества. Они являются основой для заключения юридических сделок, удостоверения личности, проведения финансовых операций и множества других важнейших процессов.

В условиях глобализации и цифровизации количество документов, с которыми ежедневно работают организации и учреждения, постоянно растёт. Одновременно с этим возрастает угроза фальсификации этих документов, что может привести к серьёзным последствиям, включая финансовые потери, нарушение прав и законных интересов граждан и организаций.

Поэтому необходимость надёжной верификации подлинности документов становится всё более актуальной.

Одним из наиболее простых и распространённых способов подтверждения подлинности документов является подпись. Подпись представляет собой уникальный графический символ, который ставит человек на документе для подтверждения своего согласия с его содержанием. Она используется с давних времен, из-за этого люди научились её подделывать, и в некоторых случаях настолько искусно, что оригинал почти не отличить от подлинника.

Из-за своей распространённости и простоты от подписи не отказались, место этого была придумана наука: графология, для нахождения связи между человеком и его подчерком. Которую потом заменили на почерковедение, изучающая методы, которые помогут найти особенности письма и почерка с целью их анализа, идентификации и проверки подлинности документов. Сам процесс проверки подлинности назвали почерковедческая экспертиза, которая рекомендовала себя, как гарантированный способ проверить подлинность подчерка.

Сама почерковедческая экспертиза состоит из анализа: визуального, гравиметрического, сравнительного, динамики написания и микроскопии. Каждый процесс занимает время и человеческие ресурсы. Необходимо провести, если не каждый, то половину из перечисленных анализов, иначе пострадают невиновные в подделывании подписи или виновники в подделывании останутся без наказания.

Этот важнейший этап требует оптимизации. Сократить время работы и требуемые человеческие ресурсы без потери точности, преобразовав, если не все, то половину аналитических методов в автоматизированный процесс. Оптимизация этого этапа позволит специалистов почерковедения сконцентрировать всё свое внимание на анализы высокой классификации, а не специалистом почерковедения (банкирам, чиновникам, полиции и т. д.), самостоятельно выявить поддельные подписи низкого и среднего качества. Тем самым ускоряя процесс без потери точности и снижения человеческих ресурсов.

Шаблонные решения для методов аналитики почерковедения для выявления подлинности подписи без участия человека не существует

Цель проекта – Разработка модели автоматической проверки подписи на основе сиамских нейросетей.

Задачами являются:

1. Поиск информации и анализ предметной области;
2. Анализ подходящих моделей;
3. Разработка плана создания модели;
4. Создание модели;
5. Подключение модели к базе данных;
6. Создание приложения для взаимодействия с моделью

# 1 Анализ предметной области и описание технологии проектирования ИС

1. 1 Описание предметной области и назначение системы

Подпись играет ключевую роль в обеспечении безопасности в идентификации человека вовремя соглашений между двух сторон. Из-за этого с давних времён присутствуют специалисты, которые ориентируются на проверке подлинности подписи.

На данный момент, подпись в современных финансовых секторах, государственных учреждениях, юридических фирмах, глобальных онлайн-платформах и частных лиц, используется для подтверждения согласия между сторонами и идентификации личности в условиях договора, а также в различных документах.

В связи с развитием технологий и доступа к информации о людях, подделка подписей становится все более доступной и трудной для обнаружения, что приводит к увеличению числа случаев мошенничества.

В свою очередь, основной проблемой традиционных методов проверки подписей являются:

* Наличие узконаправленных специалистов, которые знают методики анализа подчерков, таких как визуальный анализ, почерковедческое исследование и т. д.;
* Методы оказываются всё более временно затратными, либо более быстрыми по времени анализу подписи, но менее эффективными в обнаружении поддельной подписи.

Как итог современному миру для безопасности необходимо обеспечить всё более точными и одновременно с этим более быстрыми методами анализа подписей, которые не потребуют от исполнителя узконаправленных знаний в методиках анализа подписей.

Технология автоматической проверки подписей на основе машинного обучения появилась как решение этих проблем. Она позволяет использовать мощные алгоритмы для анализа образцов подписей и выявления даже малейших отличий между подлинной и поддельной подписью, так же благодаря автоматизации процесса верификации подписей значительно ускоряет обработку транзакций, уменьшает риски мошенничества и снижает нагрузку на специалистов.

Одним из самых перспективных методов для решения этой задачи являются сиамские нейронной сети, которые работают по принципу сравнения двух объектов, выявляя их сходства или различия. Такой подход идеально подходит для задачи распознавания и проверки подписей, так как позволяет учитывать даже малейшие различия, которые могут быть недоступны для человеческого глаза.

1.2 Целевые группы пользователей

Разработанная система для автоматической проверки подписей с использованием машинного обучения ориентирована на несколько категорий пользователей, каждый из которых сталкивается с необходимостью верификации подписи в своей профессиональной деятельности.

1. Банковские учреждения
   * Описание: Банки и финансовые организации используют подписи для аутентификации различных документов, таких как кредитные договоры, кредитные заявки, а также для подтверждения транзакций. Эти учреждения проводят большое количество операций, которые требуют проверки подлинности подписей, как для защиты от мошенничества, так и для соблюдения нормативных требований.
   * Потребности: Банковским учреждениям необходимо решение, которое позволит быстро и точно проверять подписи на документах. В условиях больших объемов операций традиционные методы проверки, такие как ручная верификация или использование простых алгоритмов, недостаточно эффективны. Это приводит к задержкам в обслуживании клиентов и увеличивает риск и критичность ошибок. Для повышения безопасности и сокращения времени на обработку операций необходимо внедрять системы, использующие машинное обучение. С помощью нейросетевых технологий можно значительно ускорить процесс проверки подписей, обеспечив их точную аутентификацию с минимальными затратами времени и ресурсов.
2. Государственные органы и учреждения
   * Описание: В государственных структурах подпись является важным атрибутом для подтверждения подлинности официальных документов, таких как налоговые декларации, контракты, решения судов и другие правовые акты. Эти документы часто имеют критическое значение для соблюдения законности и правопорядка.
   * Потребности: Государственные органы сталкиваются с необходимостью быстрой и точной проверки подлинности подписей на официальных документах, что требует значительных ресурсов и времени. Традиционные методы проверки подписей, как правило, менее эффективны в условиях большого объема документации. Ошибки в этом процессе могут привести к юридическим последствиям и затруднениям в работе. Внедрение системы на базе машинного обучения поможет ускорить процесс верификации, повысив точность и надежность, что позволит избежать ошибок и обеспечит более эффективное обслуживание.
3. Юридические фирмы и нотариальные конторы
   * Описание: Юридические фирмы и нотариальные конторы регулярно работают с документами, которые требуют обязательной проверки подписей. Эти организации играют ключевую роль в удостоверении сделок, составлении договоров и других юридически значимых актов. Подпись, как основной элемент подтверждения, должна быть проверена с максимальной точностью.
   * Потребности: Юридические специалисты сталкиваются с необходимостью проверки большого количества подписей ежедневно, что требует значительных затрат времени и усилий. Ошибки в процессе проверки могут привести к юридическим последствиям, таким как признание сделки недействительной или утрата прав на имущество. Система, использующая машинное обучение, поможет автоматизировать этот процесс, существенно сократив время на проверку и повысив точность, что особенно важно для юридических и нотариальных услуг.
4. Частные пользователи и другие лица
   * Описание: Частные пользователи сталкиваются с необходимостью подписания документов в цифровом формате, таких как договора, соглашения, налоговые декларации и другие юридические бумаги. В условиях активного перехода к цифровым услугам важно обеспечить безопасность и подлинность подписи.
   * Потребности: Частные пользователи нуждаются в удобных и надежных системах для проверки подлинности своей подписи, чтобы гарантировать её защищенность при подписании документов онлайн. В настоящее время многие пользователи сталкиваются с проблемой подделки подписей или недостаточной гарантией их безопасности в электронном виде. Для защиты от мошенничества и обеспечения доверия к подписанным документам необходимо внедрять решения на базе машинного обучения, которые обеспечат надежную верификацию подписей и защиту персональных данных, что повысит уровень безопасности и доверия со стороны пользователей.

Сиамская модель для автоматизированной проверки подписи нацелена на сферы, такие как банковские учреждения, государственные органы и учреждения, юридические фирмы и нотариальные конторы, частные пользователи и другие лица, в которых присутствуют огромное количество специалистов и даже рядовых людей. Это решение позволит всем этим сферам, соответственно и огромному количеству группам специалистов и рядовых людей сократить время на проверку подписей, при этом имея высокую точность их проверки, благодаря чему доверие к подписям в документах и т. д. не пропадёт.

1.3 Ограничения проектного решения

Ограничения проектного решения по созданию модели автоматической проверки подписи на основе сиамских нейросетей включают несколько ключевых аспектов, которые могут повлиять на процесс разработки, внедрения и эксплуатации:

1. Недостаток данных для обучения
   * Описание: Для обучения модели с использованием машинного обучения требуется достаточное количество качественных данных. В случае с проверкой подписей, нехватка разнообразных данных о подписях может ограничить эффективность модели. Это касается как подлинных подписей, так и различных типов подделок.
   * Влияние: Недостаток данных может привести к неспособности модели корректно обрабатывать все типы подписей и подделок, что снижает её точность и применение в реальных условиях.
2. Качество данных
   * Описание: Качество данных играет критическую роль в точности работы модели. Изображения подписей, которые имеют низкое разрешение или искажения, могут повлиять на способность нейросети правильно распознавать подписи.
   * Влияние: Плохое качество данных может снизить точность системы, затруднив правильную верификацию подписей и увеличив вероятность ошибок при обработке документов.
3. Ограниченная адаптивность модели
   * Описание: Модели на основе машинного обучения ограничены в своей способности адаптироваться к новым или ранее не встречавшимся типам подписей.
   * Влияние: Если система не обучена на достаточном количестве разнообразных подписей, она может не распознать новые или нестандартные подписи, из-за переобучения, что приведет к недостоверным результатам в реальных условиях.
4. Вычислительные ресурсы
   * Описание: Для обучения и работы сложных нейросетевых моделей требуется мощная вычислительная инфраструктура, включая GPU и серверы с высокой производительностью.
   * Влияние: Недостаток вычислительных ресурсов может замедлить процесс обучения и снизить скорость работы системы, что особенно критично при большом объеме данных и необходимости обработки подписей в реальном времени.
5. Возможные погрешности и ошибки
   * Описание: Даже при высокой точности системы, ошибки в верификации подписей возможны, особенно когда модель сталкивается с нестандартными подписями.
   * Влияние: Ошибки могут привести к ложным срабатываниям, что создаст проблемы для пользователей и снизит доверие к системе.
6. Необходимость в быстром и точной модели
   * Описание: При низкой скорости, можно заиметь хорошую точность, и наоборот, но модель должна быть и быстрой, и точной из-за современного мира, где огромное количество заявок и частой опасности подделывания подписи
   * Влияние: низкая скорость может привести недовольству людей и специалистов, при этом низкая точность могут привести к недоверию модели, что создаст проблемы для целевых пользователей и снизит поток заявок.

Ограничения проектного решения связаны с необходимостью большого объёма данных, максимальной точности распознавания поддельной или подлинной подписи, требованием иметь одновременно и точность, и быстроту метода проверки подписи, а также лёгкость пользования рядовым человеком или специалистом в определённой сфере, не связанной с почерковедческими науками. Эти аспекты нужно учитывать при разработке сиамской модели, чтобы она была максимально гибкой, адаптируемой, быстро, точной и удобной в использовании.

1.4 Анализ предметной области на уровне построения диаграммы использования системы (USE-CASE) и ER-диаграммы

Для эффективного проектирования и анализа системы автоматической проверки подписей на основе сиамских нейросетей необходимо визуализировать её работу через диаграммы. Для этого хорошо подходят: диаграммы использования системы (USE-CASE) и ER-диаграммы. Это позволяет описать ключевые функции системы, взаимодействие пользователей с ней и структуру данных, которые она использует.

### 1.4.1 Диаграмма вариантов использования (Use-Case)

Диаграмма вариантов использования описывает взаимодействие системы с её пользователями и основные функции, которые она предоставляет. В случае системы проверки подписей ключевыми актёрами и функционалом системы являются:

1. Пользователь, который загружает подпись для проверки и получает результаты, так же имеет возможность получить справку об системе.
2. Администратор, отвечающий за управление моделью и настройками системы для более эффективной работой системы.
3. Система проверки подписей, выполняющая анализ, создании отчётности и выдающая результаты.

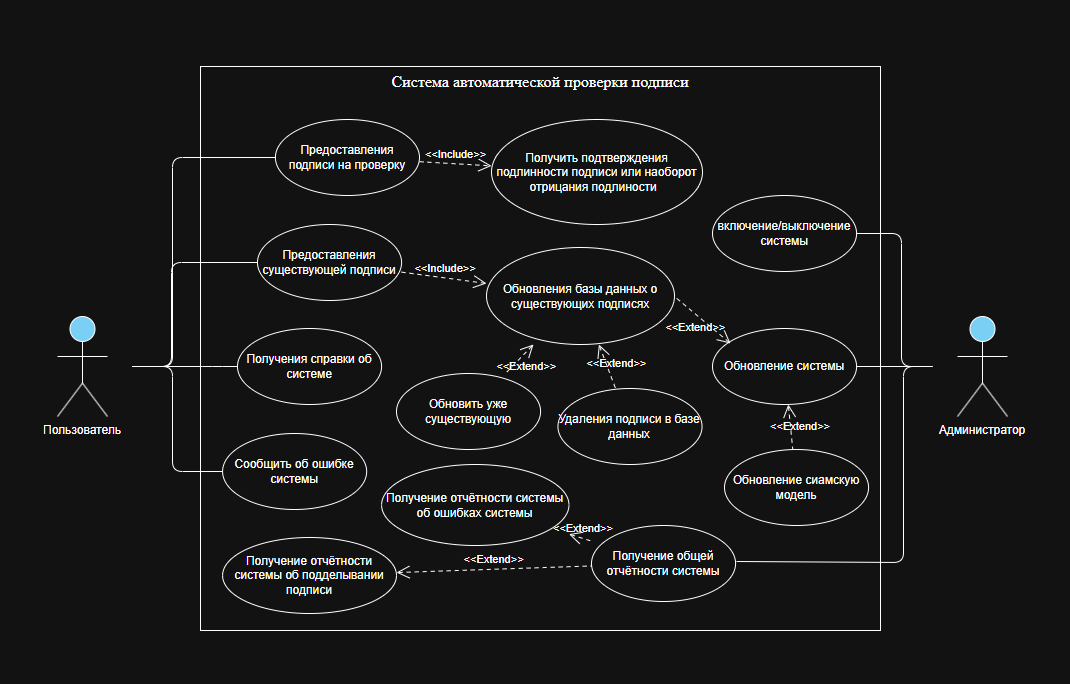
Диаграмма сценариев использования (Use-Case):

Рисунок 1 – Диаграмма сценариев использования (Use-Case)

### 1.4.2 ER-диаграмма (Диаграмма сущность-связь)

ER-диаграмма описывает ключевые данные, которые используются в системе автоматической проверки подписей, и их взаимосвязи. Эта структура помогает понять, как данные организованы, обрабатываются и хранятся, чтобы обеспечить точную и надёжную проверку подписей.

ER-диаграммы:

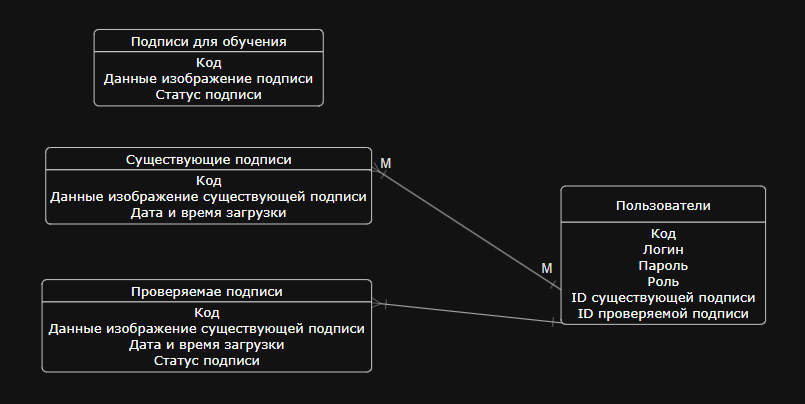
Взаимосвязи:

Рисунок 1 – ER-диаграммы

1. Один пользователь может загрузить несколько подписей.
2. Каждая подпись связана с единственным пользователем.
3. Подпись имеет один результат проверки, указывающий на её подлинность или подделку.
4. Система обновляется на основе исторических данных, которые используются для её обучения.

1.4.3 Вывод

Анализ системы автоматической проверки подписей с использованием диаграмм предоставил всестороннее понимание её архитектуры и ключевых процессов.

Диаграмма вариантов использования (Use-Case) позволила визуализировать взаимодействие пользователей с системой и администратора, выделив основные сценарии работы, такие как загрузка подписей, их проверка, просмотр результатов и управление моделью, а также проверка отчетности, включение и выключение системы и её обновления. Это помогает определить роли пользователей и администратора, их потребности и основные функции системы.

ER-диаграмма показала структуру данных и их взаимосвязь. Она помогла выделить ключевые сущности — пользователей, существующие подписи, подписи для проверки и данные для обучения. Такая структура обеспечивает чёткое понимание того, как данные обрабатываются, сохраняются и используются в процессе анализа подписей. Взаимосвязи между сущностями подчеркивают логику работы системы и облегчают её дальнейшую разработку.

Анализ с помощью диаграмм не только упрощает проектирование системы, но и помогает выявить возможные узкие места и улучшить процессы. Такой подход закладывает основу для создания надёжной, масштабируемой и функциональной системы автоматической проверки подписей.

2 Требования к системе

2.1 Требований к функциональности приложения

Система для проверки подписи должна обеспечивать следующие функциональные возможности с учётом многопользовательского режима работы:

* Обеспечение работы сиамской нейронной сети;
* Получение изображений.

2.2 Требования к графическому интерфейсу

Графический интерфейс разрабатываемого приложения должен соответствовать следующим параметрам:

* + - 1. Привлекательность – дизайн интерфейса должен быть современным и привлекательным, также соответствовать приятной цветовой гамме и легко читаемости шрифтов;
      2. Навигация – упрощённая структура навигации, для более интуитивного и понятного нахождения тех или иных функций;
      3. Гибкость – возможность быстрой обновляемости системы с добавлением новых функций и исправления ошибок.

Если следовать этим требованиям, то они помогут в создании эффективной и удобной системы проверки подписи, способной удовлетворить потребности целевой аудитории данной системы, а также поддерживать стабильность и безопасность работы приложения.

2.3 Требования к надёжности системы

Надёжность системы является одним из ключевых критериев, обеспечивающих её функциональность и безопасность. Учитывая, что каждая подпись имеет конфиденциальный статус и при проверке пользователь должен быть уверен, что его подпись не будет украдена, система должна соответствовать строгим требованиям. Это необходимо для защиты данных и предотвращения их несанкционированного доступа.

Для обеспечения безопасности системы автоматической проверки подписей необходимо предусмотреть минимизацию рисков, связанных с возможными сбоями и утечками данных. Это требует внедрения современных решений по защите информации, включая шифрование данных, мониторинг действий пользователей и управление доступом. Важной частью системы является организация надежного резервного копирования и обеспечение быстрого восстановления данных в случае чрезвычайных ситуаций, что позволит гарантировать их сохранность и доступность в любых условиях.

Все подсистемы должны пройти тестирование на устойчивость к возможным угрозам, как внешним, так и внутренним, включая попытки взлома системы, аппаратные сбои и человеческие ошибки. Использование отказоустойчивых серверов, дублирование критически важных узлов и регулярные обновления программного обеспечения являются обязательными мерами для обеспечения стабильности работы системы.

При этом, необходимо доказательства того, что информация, хранящаяся в системе, будет доступна только для уполномоченных пользователей. Для этого необходимо внедрение системы авторизации и строгих политик безопасности, что исключит несанкционированный доступ. Все данные, которые передаются или сохраняются в системе, должны быть надежно защищены от перехвата, подмены или утраты с использованием самых современных методов защиты.

В результате, надежность системы является основой её эффективной работы, а также для поддержания доверия со стороны пользователей. Поддержание высокого уровня безопасности и стабильности системы должно быть приоритетом на всех этапах её разработки, внедрения и эксплуатации.