**Министерство науки и высшего образования**

**Российской Федерации**

**ФЕДЕРАЛЬНОЕ Государственное Автономное**

**образовательное учреждение высшего образования**

**«новосибирский национальный исследовательский государственный университет»**

**Факультет информационных технологий**

Кафедра общей информатики

Направление подготовки 09.03.01 Информатика и вычислительная техника

Направленность (профиль) Программная инженерия и компьютерные науки

**ОТЧЕТ**

**о прохождении производственной практики (научно-исследовательской работы)**

(указывается наименование практики)

**Обучающегося** Кондренко Кирилла Павловича **группы №** 21202 **курса 4**

(Ф.И.О. полностью)

**Тема задания**: Проектирование серверной части системы для анализа и индивидуального подбора косметических средств

**Место прохождения практики:** ФГБУН Институт математики им. С. Л. Соболева СО РАН, Лаборатория теории вычислимости и прикладной логики, 630090, Новосибирская обл., Новосибирск, пр. Академика Коптюга, 4

(полное наименование организации и структурного подразделения, индекс, адрес)

**Сроки прохождения практики:** с 30.09.2024 г. по 23.12.2024 г.

**Руководитель практики   
от профильной организации** Пальчунов Дмитрий Евгеньевич, в.н.с.

(Ф.И.О. полностью, должность) (подпись)

**Руководитель практики от НГУ** Яхъяева Гульнара Эркиновна, доцент

(Ф.И.О. полностью, должность) (подпись)

**Руководитель ВКР** Ряскин Александр Николаевич доцент

(Ф.И.О. полностью) (должность)

**Оценка по итогам защиты отчета:**

(неудовлетворительно, удовлетворительно, хорошо, отлично)

**Отчет заслушан на заседании кафедры** общей информатики

(наименование кафедры)

**протокол \_\_\_\_\_\_\_\_\_от** «\_\_\_\_\_\_» декабря 2024 г.

Новосибирск 2024

**СОДЕРЖАНИЕ**

[Введение 3](#_Toc185533753)

[Основная часть 5](#_Toc185533754)

[1. Общие сведения о серверной части 5](#_Toc185533755)

[2. Алгоритм рекомендации косметических средств 5](#_Toc185533756)

[2.1. Теоретические сведения 5](#_Toc185533757)

[2.2 Разработанный алгоритм 5](#_Toc185533758)

[3. Программные модули 6](#_Toc185533759)

[4. Аутентификация и авторизация пользователей 7](#_Toc185533760)

[5. Базы данных, языки программирования, фреймворки, библиотеки 7](#_Toc185533761)

[Заключение 8](#_Toc185533762)

[Список использованных источников и литературы 9](#_Toc185533763)

[Приложение А 10](#_Toc185533764)

[Приложение Б 11](#_Toc185533765)

# ВВЕДЕНИЕ

В современном обществе наблюдается значительное увеличение ассортимента косметических средств и средств для ухода за собой. На рынке косметической продукции складывается острая конкуренция между производителями, продавцами и поставщиками, каждый из которых стремится привлечь внимание потребителей и добиться успеха за счет эффективных рекламных кампаний. Однако зачастую доверие к продуктам крупных брендов основывается не только на их реальной эффективности и безопасности, но и на популярности самой марки, что приводит к определенным перекосам в восприятии потребителей.

Одновременно с этим растут и требования потребителей. Современные пользователи ожидают от косметических средств не только соответствия высоким стандартам безопасности и использования натуральных компонентов, но и мгновенного видимого эффекта. Важным критерием также остается соотношение цены и качества, что делает выбор подходящего продукта еще более сложной задачей.

С развитием цифровых технологий многие аспекты повседневной жизни стали более простыми и удобными. Значительный вклад в это внесло появление индивидуальных помощников и систем автоматизации, способных анализировать данные и предлагать оптимальные решения на основании заданных пользователем параметров.

**Целью** практики являлась проектирование серверной части индивидуального помощника, который сможет выполнить роль виртуального консультанта при подборе косметических средств для ухода за собой.

Для достижения поставленной цели были поставлены следующие **задачи**:

1. разработка алгоритма рекомендаций косметических средств,
2. определение программных модулей серверной части,
3. выбор/разработка механизма аутентификации и авторизации пользователей,
4. выбор используемых баз данных,
5. выбор языков программирования, фреймворков и библиотек для дальнейшей разработки.

Также дополнительными задачами на практику являлись:

1. выступление с докладом на конференции «Мальцевские чтения — 2024» (приложение А),
2. написание текста введения ВКР (приложение Б).

Актуальность темы практики обусловлена тем, что у многих людей часто возникает потребность в выборе средств по уходу за собой, однако существует проблема, которая состоит в том, что не все разбираются в косметических средствах настолько хорошо, чтобы подобрать себе правильную линейку средств и оценить их состав при покупке и подборе. Это приводит к неэффективным покупкам, расходам времени, а также может повлечь применение опасных для здоровья косметических средств ввиду индивидуальных особенностей организма, таких как аллергия на определенные компоненты в составе косметических средств, например. Проектирование серверной части системы для анализа и индивидуального подбора косметических средств позволит создать приложение, которое решит обозначенную проблему.

В рамках практики объектом исследования являлись рекомендательные системы, а предметом исследования — алгоритм подбора косметических средств, учитывающий поведение и предпочтения пользователей.

**Сведения о месте прохождения практики:** Место прохождения практики: ФГБУН Институт математики им. С. Л. Соболева СО РАН, Лаборатория теории вычислимости и прикладной логики, 630090, Новосибирская обл., Новосибирск, пр. Академика Коптюга, 4. Главная цель Института состоит в создании фундаментальных знаний в развиваемых им научных направлениях, разработке научных основ современной техники и технологии без цели извлечения прибыли.

# ОСНОВНАЯ ЧАСТЬ

## 1. Общие сведения о серверной части

Программа должна представлять собой веб-сервер, предоставляющий API, который совместим с RestAPI. То есть входные и выходные данные должны быть организованы, то есть приниматься и отдаваться, в формате JSON по протоколу HTTP; будут использованы стандартные единицы измерения дат, валют и времени.

## 2. Алгоритм рекомендации косметических средств

### 2.1. Теоретические сведения

Все теоретические сведения о рекомендательных системах были взяты из источника [1].

**Общий подход рекомендательных систем**:

* *U* – множество субъектов (пользователей);
* *I* – множество объектов (косметические средства);
* *Y* – описание транзакций (действий субъектов над объектами);
* *R* = (*rui*) – матрица, где *rui* – насколько субъекту *u* «нравится» объект *i*.

**Задачи рекомендательных систем**:

1. восстановление пропущенных значений в матрице *R*;
2. построение множества близких субъектов и объектов по матрице *R*.

**Типы моделей рекомендательных систем**:

1. Корреляционные модели. Данных подход использует всю матрицу R и судит о сходствах объектов и субъектов по корреляции строк и столбцов матрицы.
2. Латентные модели. В данной подходе объектам и субъектам сопоставляются векторные представления, по близости которых оценивается близость субъектов и объектов.

### 2.2 Разработанный алгоритм

Подход, который используется в разработанном алгоритме, относится к латентным моделям. Данная модель выбрана, поскольку корреляционная модель имеет следующие недостатки:

1. имеет место проблема “Холодного старта”, то есть недавно добавленные косметические средства не будут никому рекомендоваться;
2. требуется хранение всей матрицы R;

в то время как латентная модель лишена этих недостатков.

Для подбора косметических средств используются следующие индивидуальные особенности пользователей:

1. Пол.
2. Возраст.
3. Область применения (волосы, тело, губы, лицо и т.д.). Для корректной работы необходимо выбрать хотя бы одну область применения.
4. Назначение (восстановление кожи губ, увлажнение кожи лица, акне и чёрные точки на лице и т.д.). Выбирается по области применения.
5. Предпочитаемый тип продукта (шампунь-гель, масло, крем, бальзам, кондиционер и т.д.). Это необязательный параметр.
6. Дополнительные особенности (жирная кожа, сухая кожа, ломкие волосы, окрашенные волосы и т.д.). Выбираемые особенности должны соответствовать выбранным назначениям.

Перед началом работы алгоритмы всем косметическим средствам нужно сопоставить векторное представление, соответствующее индивидуальным особенностям пользователей. Для этого можно, например, использовать ключевые слова в названиях и описаниях косметических средств. После того, как косметическому средства был сопоставлен вектор, он может рекомендоваться пользователям.

Последовательность действий для подбора косметических средств:

1. Пользователь регистрируется в системе, указывая данные, которые не влияют на подбор косметических средств;
2. При подборе пользователь указывает свои индивидуальные особенности, которым сопоставляется векторное представление *v*;
3. Происходит обращение к векторной базе данных с запросом по поиску *N* косметических средств, векторные представления которых наиболее близки к *v*.
4. Полученные от векторной базы данных косметические средства рекомендуются пользователь в порядке возрастания векторного расстояния до *v*.

## 3. Программные модули

Серверная часть будет состоять из следующих программных модулей:

1. модуль, отвечающий за анализ состава косметических средств посредством обращения к базе данных составов и ингредиентов косметических средств cosmobase.ru [2];
2. модуль, отвечающий за рекомендации косметических средств, который будет строить векторные представления и делать запросы к векторной базе данных;
3. модуль, отвечающий за хранение данных о пользователях, которые не связаны с подбором косметических средств; он будет обращаться к реляционной базе данных.

## 4. Аутентификация и авторизация пользователей

Аутентификацию и авторизацию будут основаны на механизме OAuth2.0 [3] с использованием JWT [4]. Как следствие, должны использоваться два токена: обновления и доступа. Данные о пользователях будут храниться в реляционной базе данных, при этом пароли пользователей хранятся не в открытом в виде, а в виде хэша. В качестве функции хэширования была выбрана функция «Argon2» [5, 6].

## 5. Базы данных, языки программирования, фреймворки, библиотеки

в качестве языка разработки был выбран Python 3.12 [7], а в качестве фреймворка – FastAPI 0.115.x [8];

Хранение данных должно обеспечиваться с использование следующих баз данных:

1. Weaviate [9] — векторная база данных, в которой предполагается хранить векторные представления пользователей и косметических средств. Должна быть использована в виде docker-контейнера из образа cr.weaviate.io/semitechnologies/weaviate:1.27.6.
2. PostgreSQL [10] — реляционная база данных, в которой предполагается хранить информацию о пользователях. Должна быть использована в виде docker-контейнера из образа postgres:17.2.

При разработке будут использоваться следующие сторонние библиотеки:

1. fastapi 0.115.x — создание API на основе RestAPI;
2. aiohttp 3.x — асинхронная отправка HTTP-запросов;
3. pyjwt 2.x — выпуск и проверка JWT;
4. structlog 24.x — логирование работы серверного модуля;
5. argon2-cffi — хеширование паролей пользователей с использованием алгоритма Argon2;
6. fastapi-mail 1.4.x — для отправки электронной почты (взаимодействие с  серверами, поддерживающими протоколы электронной почты);
7. sqlalchemy 2.x — взаимодействие с PostgreSQL;
8. weaviate-client 4.x — взаимодействие с Weaviate.

Для управления используемыми библиотеками будет использован инструмент Poetry 1.8.x.

# ЗАКЛЮЧЕНИЕ

В результате прохождения практики были выполнены следующие задачи:

1. разработан алгоритм рекомендаций косметических средств, основанный на векторном представлении пользователей и косметических средств;
2. определены программных модулей серверной части;
3. выбран механизм аутентификации и авторизации пользователей, основанный на двух видах JWT: токены доступа и токены обновления;
4. в качестве баз данных серверной части были выбраны PostgreSQL 17.x

и Weaviate 1.27.x;

1. в качестве языка разработки был выбран Python 3.12, а в качестве фреймворка – FastAPI 0.115.x;
2. написан текст введения ВКР,
3. по теме практики было выступление с докладом на конференции «Мальцевские чтения — 2024».

Таким образом, была спроектирована серверной части системы для анализа и индивидуального подбора косметических средств, которая может быть использована для создания приложения, которое будет включать в себя как клиентскую, так и серверную части.

В ходе прохождения практики были изучены механизмы аутентификации и авторизации пользователей, применяемые в бэкенд-разработке, а также принципы работы рекомендательных систем; был получен опыт в проектировании бэкенд-приложений.

# СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ И ЛИТЕРАТУРЫ

1. Машинное обучение: курс лекций К. В. Воронцов: Коллаборативная фильтрация и матричные разложения. URL: http://www.machinelearning.ru/wiki/images/9/95/Voron-ML-CF.pdf (дата обращения: 19.12.2024).
2. Косметическая база данных COSMOBASE.ru. URL: https://cosmobase.ru/handbook (дата обращения: 19.12.2024).
3. RFC 6749 — The OAuth 2.0 Authorization Framework.
4. RFC 7519 — JSON Web Token (JWT).
5. RFC 9106 — Argon2 Memory-Hard Function for Password Hashing and Proof-of-Work Applications.
6. Github.com: P-H-C/phc-winner-argon2: The password hash Argon2, winner of PHC. URL: https://github.com/P-H-C/phc-winner-argon2 (дата обращения: 19.12.2024).
7. Welcome to Python.org. URL: https://www.python.org/ (дата обращения: 19.12.2024).
8. Github.com: fastapi/fastapi: FastAPI framework, high performance, easy to learn, fast to code, ready for production. URL: https://github.com/fastapi/fastapi (дата обращения: 19.12.2024).
9. The AI-native database developers love | Weaviate. URL: https://weaviate.io/ (дата обращения: 19.12.2024).
10. PostgreSQL: The world's most advanced open source database. URL: https://www.postgresql.org/ (дата обращения: 19.12.2024).

# ПРИЛОЖЕНИЕ А

(обязательное)

**ТЕЗИСЫ НА КОНФЕРЕНЦИЮ «Мальцевские чтения — 2024»**

**РАЗРАБОТКА ИНДИВИДУАЛЬНОГО ПОМОЩНИКА ДЛЯ ПОДБОРА КОСМЕТИЧЕСКИХ СРЕДСТВ**

К. В. Егоренко, К. П. Кондренко

У многих людей возникают проблемы с выбором средств для ухода за собой. Не все разбираются в косметических средствах настолько хорошо, чтобы подобрать себе правильную линейку средств еще и оценить их состав при покупке и подборе. Использование семантического подхода при разработке системы подбора косметических средств, позволяет работать с неточной и неполной информацией, учитывать индивидуальные предпочтения пользователей, а также учитывать оценочные знания пользователей о товарах [1,2].

Целью проекта является разработка индивидуального помощника, который сможет выполнить роль консультанта при подборе средств по уходу за собой (по запросу пользователя подберет все необходимые средства или линейку средств). Пользователь может просто ввести то, что ему хотелось бы подкорректировать в себе, а индивидуальный помощник предложит ему наиболее подходящие варианты линеек продуктов, вместе с этим пользователь также получит анализ состава каждого предложенного продукта и список проверенных поставщиков, у которых можно приобрести продукт.

На данный момент была реализована часть системы, отвечающая за анализ состава косметических средств и часть системы, отвечающая за аутентификацию пользователей. Сейчас ведется работа над реализацией индивидуального подбора косметических средств по заданным пользователями параметрам.

*Список литературы*

1. Yakhyaeva G.E. Application of Boolean Valued and Fuzzy Model Theory for Knowledge Base Development. SIBIRCON 2019 - International Multi-Conference on Engineering, Computer and Information Sciences, Proceedings, 2019, pp. 868–871.
2. Yakhyaeva G., Skokova V. Subjective Expert Evaluations in the Model-Theoretic Representation of Object Domain Knowledge. Lecture Notes in Computer Science, 2021, 12948 LNAI, pp. 152–165.

Новосибирский государственный университет, Новосибирск

*Email address: k.egorenko@g.nsu.ru, k.kondrenko@g.nsu.ru*

# ПРИЛОЖЕНИЕ Б

(обязательное)

**ТЕКСТ ВВЕДЕНИЯ ВКР**

**ВВЕДЕНИЕ**

**Актуальность** данной работы обусловлена тем, что у многих людей часто возникает потребность в выборе средств по уходу за собой, однако существует **проблема**, которая состоит в том, что не все разбираются в косметических средствах настолько хорошо, чтобы подобрать себе правильную линейку средств и оценить их состав при покупке и подборе. Это приводит к неэффективным покупкам, расходам времени, а также может повлечь применение опасных для здоровья косметических средств ввиду индивидуальных особенностей организма, таких как аллергия на определенные компоненты в составе косметических средств, например.

Разрабатываемый помощник призван решить данную проблему. Пользователи смогут указать свои особенности и потребности, а индивидуальный помощник предложит им наиболее подходящие варианты продуктов и линеек.

**Целью** данной выпускной квалификационной работы является разработка серверной части индивидуального помощника для подбора косметических средств, который автоматизирует процесс подбора косметики по индивидуальным особенностям пользователей и помогает анализировать составы косметических средств.

Для достижения данной цели были поставлены следующие **задачи**:

1. проанализировать текущие решения и технологий в области автоматизации подбора и рекомендательных систем;
2. составить список требований к разрабатываемому помощнику;
3. разработать и реализовать алгоритм подбора косметических средств, учитывающий индивидуальные особенности и предпочтения пользователей;
4. разработать и реализовать алгоритм анализа состава косметических средств;
5. разработать архитектуру серверной части;
6. реализовать серверную часть помощника;
7. контейнеризовать серверную часть в Docker для удобства дальнейшей разработки клиентской части;
8. провести тестирование и оценку эффективности системы на реальных пользователях.

**Научная новизна** работы заключается в разработке и реализации алгоритма рекомендаций косметических средств, который использует векторные представления пользователей и косметических средств. Для подбора используются следующие индивидуальные особенности пользователей:

1. Пол.
2. Возраст.
3. Область применения (волосы, тело, губы, лицо и т.д.). Для корректной работы необходимо выбрать хотя бы одну область применения.
4. Назначение (восстановление кожи губ, увлажнение кожи лица, акне и чёрные точки на лице и т.д.). Выбирается по области применения.
5. Предпочитаемый тип продукта (шампунь-гель, масло, крем, бальзам, кондиционер и т.д.). Это необязательный параметр.
6. Дополнительные особенности (жирная кожа, сухая кожа, ломкие волосы, окрашенные волосы и т.д.). Выбираемые особенности должны соответствовать выбранным назначениям.

**Практическая ценность** данной работы заключается в созданиисерверного модуля индивидуального помощника, который позволяет подбирать косметические средства по индивидуальным особенностям, а также поддерживает анализ состава косметических средств, аутентификацию и авторизацию пользователей. Данный модуль может быть использован для создания приложения, которое будет включать в себя и серверную, и клиентскую части. Приложение позволит пользователям экономить время при подборе косметических средств, а также исключит потенциальные ошибки пользователей, поскольку подбор будет происходить автоматически по указанным индивидуальным особенностям. Приложение будет полезно как для тех, кто хочет подобрать себе косметические средства, так и для косметологов, поскольку позволит им эффективнее оказывать косметологические услуги, ориентируясь на рекомендации индивидуального помощника.

Работа состоит из введения, трёх глав и заключения. В первой главе определяется предметная область, выполняется анализ текущих решений и технологий в области автоматизации подбора и рекомендательных систем, а также формируются требования к разрабатываемой системе. Во второй главе описывается процесс разработки: создание механизма сбора данных о пользователе, алгоритмов подбора косметических средств и анализа их составов. Третья глава посвящена тестированию разработанных алгоритмов и системы, а также их оценке на основе реальных данных. В заключении подводятся итоги работы и формулируются рекомендации для дальнейшего развития проекта.