

Урунов А.Т.

*Студент кафедры прикладной информатики
и информационных технологий*

НИУ «БелГУ», 4 курс (Белгород, Россия)

Научный руководитель: Зайцева Т.В.

*доцент кафедры прикладной информатики
и информационных технологий*

НИУ «БелГУ», (Белгород, Россия)

Urunov A.T.

*Student of the Department of Applied Informatics
and Information Technology*

NRU "BelSU", 4rd year (Belgorod, Russia)

Scientific supervisor: Zaitseva T.V.

*docent of the Department of Applied Informatics
and Information Technology*

NRU "BelGU", (Belgorod, Russia)

**АНАЛИЗ СХОДСТВА ПРЕМИАЛЬНЫХ СМАРТФОНОВ С
ИСПОЛЬЗОВАНИЕМ МЕТРИК РАССТОЯНИЯ И МЕР БИНАРНОГО
СХОДСТВА**

**ANALYSIS OF PREMIUM SMARTPHONE SIMILARITY USING
DISTANCE METRICS AND BINARY SIMILARITY MEASURES**

Аннотация: В статье рассматривается разработка программного решения для анализа сходства премиальных смартфонов на основе бинарных признаков. Для сравнения используются метрики расстояния (Евклидово, Минковского) и функции сходства (Жаккара, Дайса, Сокала–Мишнера и др.). Реализация выполнена на языке С#. Продемонстрированы результаты сравнения флагманских моделей смартфонов.

Abstract: The article presents the development of a software solution for analyzing the similarity of premium smartphones based on binary features. Distance metrics (Euclidean, Minkowski) and similarity measures (Jaccard, Dice, Sokal–Michener, etc.) are used for comparison. The implementation is written in C#. The results of comparing flagship smartphone models are demonstrated.

Ключевые слова: интеллектуальные информационные системы, бинарные признаки, метрики расстояния, коэффициенты сходства, анализ данных, C#.

Keywords: intelligent information systems, binary features, distance metrics, similarity coefficients, data analysis, C#.

В задачах интеллектуальных информационных систем важное место занимает анализ сходства объектов по набору признаков. Подобные методы широко применяются в системах рекомендаций, кластеризации, классификации и сравнительном анализе продукции.

В рамках лабораторной работы был реализован программный модуль для оценки степени сходства премиальных смартфонов по набору бинарных характеристик (наличие/отсутствие функции).

Для анализа были выбраны следующие флагманские модели:

- iPhone 16 Pro Max
- Samsung Galaxy S24 Ultra
- Google Pixel 9 Pro

Сравнение проводилось по следующим признакам:

1. Профессиональная камера.
2. Поддержка стилуса.
3. Сверхбыстрая зарядка.
4. Большой экран (6.7"+).
5. Продвинутая ИИ-обработка фото.

Каждый признак кодировался бинарно:

1 — функция присутствует;

0 — функция отсутствует.

Матрица признаков представлена в таблице 1.

Таблица 1 – Матрица признаков

Смартфон	Камера	Стилус	Быстрая зарядка	Большой экран	ИИ-фото
iPhone 16 Pro Max	1	0	0	1	1
Samsung Galaxy S24 Ultra	1	1	1	1	1
Google Pixel 9 Pro	1	0	1	0	1

Для количественной оценки сходства были использованы следующие метрики расстояния:

- Евклидово расстояние
- Расстояние Минковского ($p = 3$)
- Квадрат Евклидова расстояния (как расстояние n -мерных векторов)

Евклидово расстояние определяется формулой:

$$d(A, B) = \sqrt{\sum_{i=1}^n (A_i - B_i)^2}$$

Расстояние Минковского вычисляется как:

$$d(A, B) = \left(\sum_{i=1}^n |A_i - B_i|^p \right)^{1/p}$$

Кроме метрик расстояния были реализованы функции бинарного сходства:

- Рассела и Рао;
- Жаккара (Jaccard);
- Дайса (Dice);
- Сокала и Снифа;
- Сокала и Мишнера;
- Кульжинского;
- Сокала и Юла.

Пусть:

a — количество совпадений по признаку 1–1;

b — совпадения 0–0;

g — 1–0;

h — 0–1.

Тогда, например, коэффициент Жаккара определяется как:

$$J = \frac{a}{a + g + h}$$

Программа реализована на языке C# в виде консольного приложения. В ней заданы массивы признаков для каждого смартфона, после чего выполняется попарное сравнение моделей. Расчёт расстояний и коэффициентов сходства осуществляется в отдельных функциях, что делает архитектуру программы модульной и удобной для расширения.

Результаты вычислений показали, что наибольшая степень сходства наблюдается между моделями Samsung Galaxy S24 Ultra и Google Pixel 9 Pro,

поскольку они имеют совпадение по большинству функциональных характеристик. Наименьшее сходство выявлено между iPhone 16 Pro Max и Google Pixel 9 Pro из-за различий по размеру экрана и поддержке быстрой зарядки.

Разработанное программное решение демонстрирует применение математических методов анализа сходства в задачах интеллектуальных информационных систем. Подобный подход может быть расширен для многомерных данных, весовых коэффициентов признаков и интеграции в рекомендательные системы электронной коммерции.

Таким образом, в работе была реализована и протестирована система сравнительного анализа объектов на основе бинарных признаков с использованием различных метрик расстояния и коэффициентов сходства, что подтверждает эффективность математического аппарата теории сходства в практических задачах анализа данных.

Использованные источники:

1. Хастие, Т. Основы статистического обучения: интеллектуальный анализ данных, вывод и прогнозирование / Т. Хастие, Р. Тибширани, Дж. Фридман. – Москва: ДМК Пресс, 2017. – 768 с.
2. Айвазян, С. А. Прикладная статистика: классификация и снижение размерности / С. А. Айвазян, В. М. Бухштабер, И. С. Енюков. – Москва: Финансы и статистика, 1989. – 607 с.
3. Тан, П.-Н. Введение в анализ данных и машинное обучение / П.-Н. Тан, М. Стайнбах, В. Кумар. – Москва: Вильямс, 2016. – 864 с.
4. Трофимов, В. В. Язык программирования C# и платформа .NET / В. В. Трофимов. – Санкт-Петербург: Питер, 2022. – 640 с.