

Журнал зарегистрирован Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций (Роскомнадзор)

Свидетельство о регистрации
ПИ № ФС77-64098
от 18 декабря 2015 г.

Учредитель – Федеральное государственное бюджетное образовательное учреждение высшего образования «Технологический университет имени дважды Героя Советского Союза, летчика-космонавта А.А. Леонова» (141074, Московская область, г. Королев, ул. Гагарина, д. 42)
Издается с сентября 2014 г.
Выходит 4 раза в год

ISSN 2409-1650

Журнал «Информационно-технологический вестник» включён в Перечень ведущих периодических изданий ВАК

Группы научных специальностей и научные специальности в рамках групп научных специальностей, по которым издание входит в Перечень:
2. Технические науки; 2.2. Электроника, фотоника, приборостроение и связь; 2.3. Информационные технологии и телекоммуникации [2.3.1. Системный анализ, управление и обработка информации; 2.3.5. Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей]; 2.5. Машиностроение [2.5.13. Проектирование, конструкция и производство летательных аппаратов]; 2.6. Химические технологии, науки о материалах, металлургия; [2.6.17. Материаловедение]

Подписной индекс в каталоге «Почта России» ПП1997

Главный редактор
Артюшенко Владимир Михайлович,
д.т.н., профессор

Над выпуском работали
Паршина Ю.С.
Пирогова Е.В.

Адрес редакции:
141070, Королев,
Ул. Октябрьская, 10а
Тел. (495)543-34-30 (доб.138),
E-mail: rio-kimes@mail.ru,
Site: www.unitech-mo.ru

Редакция не несет ответственности за достоверность информации в материалах, в том числе рекламных, представленных авторами для публикации
Материалы приводятся в авторской редакции.

СОДЕРЖАНИЕ

СИСТЕМНЫЙ АНАЛИЗ, УПРАВЛЕНИЕ И ОБРАБОТКА ИНФОРМАЦИИ

Герасимов В.А., Шайтура С.В.
СИСТЕМА ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ДЛЯ ОПТИМИЗАЦИИ ИНФОРМАЦИОННЫХ СИСТЕМ.....3

Истратова Е.Е., Кожевников А.Н.
РАЗРАБОТКА ЦИФРОВОГО ДВОЙНИКА ИНФОРМАЦИОННОЙ СИСТЕМЫ ПРЕДПРИЯТИЯ ДЛЯ ВНЕДРЕНИЯ В ОБРАЗОВАТЕЛЬНУЮ ПРОГРАММУ ВЫСШЕГО УЧЕБНОГО ЗАВЕДЕНИЯ.....14

Шульженко С.Н., Перепелица К.А.
МЕТОД УПРАВЛЕНИЯ РИСКАМИ В БОЛЬШИХ ИНФОРМАЦИОННЫХ СИСТЕМАХ.....22

МАТЕМАТИЧЕСКОЕ И ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ, КОМПЛЕКСОВ И КОМПЬЮТЕРНЫХ СЕТЕЙ

Ковалев И.И., Аббасова Т.С.
ПОДДЕРЖКА ПРИНЯТИЯ РЕШЕНИЙ НА ОСНОВЕ ПРАВИЛ ДОКАЗАТЕЛЬНОЙ АРГУМЕНТАЦИИ.....34

Хорев П.Б., Павлов И.С.
ЭФФЕКТИВНАЯ РЕАЛИЗАЦИЯ КРИПТОАЛГОРИТМОВ РОССИЙСКИХ СТАНДАРТОВ НА МОБИЛЬНЫХ УСТРОЙСТВАХ.....48

Шайтура Н.С., Белова Ю.Н., Феоктистов Е.А.
ГЕОИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ТОЧНОГО ЗЕМЛЕДЕЛИЯ.....64

ПРОЕКТИРОВАНИЕ, КОНСТРУКЦИЯ И ПРОИЗВОДСТВО ЛЕТАТЕЛЬНЫХ АППАРАТОВ

Артюшенко В.М., Семенов А.Б.
МАТЕМАТИЧЕСКОЕ ОБЕСПЕЧЕНИЕ МОДЕЛИРОВАНИЯ ЭЛЕКТРОМАГНИТНЫХ ВОЗМУЩЕНИЙ НА ПОВЕРХНОСТИ ТЕХНИЧЕСКИХ ОБЪЕКТОВ В ПОТОКЕ ИОНИЗИРУЮЩИХ ИЗЛУЧЕНИЙ.....83

Бершадский В.А.
ОПРЕДЕЛЕНИЕ ЭФФЕКТИВНОСТИ ПРОЦЕССА НАДУВА ТОПЛИВНЫХ БАКОВ РАКЕТНОЙ ДВИГАТЕЛЬНОЙ УСТАНОВКИ.....94

Вокин Г.Г., Насибулин М.Ш.
КОНЦЕПТУАЛЬНЫЕ ОСНОВЫ МАКСИМИЗАЦИИ ЭФФЕКТИВНОСТИ ПЕРСПЕКТИВНЫХ СИСТЕМ УПРАВЛЕНИЯ ВАЖНЕЙШИХ РАКЕТНО-КОСМИЧЕСКИХ ЛЕТАТЕЛЬНЫХ АППАРАТОВ НА ВСЕХ ЭТАПАХ ИХ ЖИЗНЕННОГО ЦИКЛА.....102

РЕДАКЦИОННЫЙ СОВЕТ

1. **Барканов Е.Н.**, Dr.sc.ing.
2. **Васильев Н.А.**, д.т.н., профессор
3. **Леоненко Д.В.**, д.ф.-м.н., профессор
4. **Тимофеев А.Н.**, д.т.н., профессор

РЕДАКЦИОННАЯ КОЛЛЕГИЯ

1. **Аббасов Э.М.**, к.т.н.
2. **Аббасова Т.С.**, к.т.н., доцент
3. **Бухаров С.В.**, д.т.н., профессор
4. **Бершадский В.А.**, д.т.н., профессор кафедры
5. **Воловач В.И.**, д.т.н., профессор
6. **Кучеров Б.А.**, к.т.н.
7. **Логачев И.А.**, к.т.н.
8. **Логачева А.И.**, д.т.н., профессор
9. **Макаров М.И.**, д.т.н., профессор
10. **Матвиенко Ю.Г.**, д.т.н., профессор
11. **Мороз А.П.**, д.т.н., профессор
12. **Мосалов О.П.**, к.ф.-м.н.
13. **Разумовский И.М.**, д.ф.-м.н., профессор
14. **Рудаков В.Б.**, д.т.н., профессор
15. **Самаров Е.К.**, д.т.н., доцент
16. **Скрябин М.Л.**, к.т.н.
17. **Соляной В.Н.**, к.т.н.
18. **Стреналюк Ю.В.**, д.т.н., профессор
19. **Халиулин В.И.**, д.т.н., профессор
20. **Чесноков А.В.**, д.т.н.
21. **Щурин К.В.**, д.т.н., профессор

Подписано в печать 25.03.2024

Формат В5

Печать офсетная. Усл.печ.л. 13,0

Тираж 500 экз.

Заказ № 97-01

Отпечатано в типографии

ООО «Научный консультант»

г. Москва

Хорошевское шоссе, 35, корп. 2

Лобанов И.Е.

**ТЕОРИЯ ТЕПЛООБМЕНА В КАНАЛАХ
С ДИАФРАГМАМИ В ЛАМИНАРНОЙ
И ПЕРЕХОДНОЙ ОБЛАСТЯХ ПРИ ТЕЧЕНИИ
ТРАНСФОРМАТОРНЫХ МАСЕЛ ДЛЯ
ПЕРСПЕКТИВНЫХ ТЕПЛООБМЕННИКОВ
ЛЕТАТЕЛЬНЫХ АППАРАТОВ.....115**

Стреналюк Ю.В., Соколова Д.А.

**ИССЛЕДОВАНИЕ МЕТОДИКИ ПРИМЕНЕНИЯ
ТЕХНОЛОГИИ ВИРТУАЛЬНОГО ОБЛАЧНОГО
ОФИСА ДЛЯ ПРОГРАММНО-АППАРАТНОЙ
ПОДДЕРЖКИ СОЗДАНИЯ ИЗМЕРИТЕЛЬНЫХ
ПРИБОРОВ СУ БЛА.....132**

МАТЕРИАЛОВЕДЕНИЕ

Абразумов В.В., Фролов С.В., Токарева О.В.

**ИЗУЧЕНИЕ ВОЗМОЖНОСТЕЙ ПРИМЕНЕНИЯ
КОМПОЗИЦИОННЫХ ПОКРЫТИЙ
С ИНТЕРКАЛИРОВАННЫМ ГРАФИТОМ
НА КРЕМНЕЗЕМНОЙ ОСНОВЕ ДЛЯ СРЕДСТВ
ЗАЩИТЫ ПРИ ПОЖАРАХ.....144**

Горлов Е.С., Пашковский И.Э.

**ОЦЕНКА ПРИМЕНИМОСТИ ТЕХНОЛОГИЧЕСКИХ
МЕТОДОВ ПОВЫШЕНИЯ СРОКА СЛУЖБЫ ДЕТАЛЕЙ
ЗА СЧЕТ СНИЖЕНИЯ ВОДОРОДНОГО
ИЗНАШИВАНИЯ.....158**

Скрябин М.Л.

**ТЕХНОЛОГИЧЕСКИЕ ОСОБЕННОСТИ
ИЗГОТОВЛЕНИЯ ОТЛИВОК ИЗ СТАЛИ 35ХГСЛ
И МЕТОДЫ ВОЗДЕЙСТВИЯ НА ПРОЦЕСС
КРИСТАЛЛИЗАЦИИ.....168**

Соловьев М.Е., Кокарев С.С., Балдаев С.Л., Балдаев Л.Х., Федорова М.О.

**РАСЧЕТ НАПРЯЖЕНИЙ НА ЛОПАСТИ
В ПОТОКЕ СРЕДЫ МЕТОДОМ КОМПЛЕКСНОГО
ПОТЕНЦИАЛА.....176**

УДК 004.05

Поддержка принятия решений на основе правил доказательной аргументации

Игорь Иванович Ковалев, кандидат технических наук, доцент,
Татьяна Сергеевна Аббасова, кандидат технических наук, доцент,
кафедра информационных технологий и управляющих систем,
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Технологический университет имени дважды Героя Советского Союза,
летчика-космонавта А.А. Леонова», г. Королев, Московская область

Проблемы современных систем поддержки принятия решений состоят в том, что немногие из них решают задачу обнаружения неточности классификации, когда в описании запроса содержится неточная или непоследовательная информация. На основе анализа числовых и текстовых характеристик запросов, а также правил доказательной аргументации, разработан классификатор запросов. Проведена оценка производительности классификации запросов, определена вероятность и степень уверенности признака запроса.

Точность запроса; чувствительность запроса; прецизионность запроса.

Decision support based on rules of evidence-based reasoning

Igor Ivanovich Kovalev, Candidate of technical Sciences, Associate Professor,
Tatyana Sergeevna Abbasova, Candidate of technical Sciences, Associate Professor,
Department of Information technologies and control systems,
Federal State Budgetary Educational Institution of Higher Education
«Technological University named after twice Hero of the Soviet Union,
pilot-cosmonaut A.A. Leonov», Korolev, Moscow region

The problem with current decision support systems is that few of them solve the problem of detecting misclassification when the query description contains inaccurate or inconsistent information. Based on the analysis of numerical and textual characteristics of queries, as well as the rules of evidentiary argumentation, a query classifier was developed. The performance of query classification was assessed and the probability and degree of confidence of the query attribute was determined.

Query accuracy; request sensitivity; query precision.

Введение

Из-за сложности телекоммуникационных технологий, когда клиенты жалуются на сбой в обслуживании или программном продукте, появляется много неопределенной или противоречивой информации [2, С.15;3, С.6;4, С.14]. Непо-

следовательная информация в описаниях заявок пользователей сильно влияет на технических специалистов, когда они выносят суждение о причинах рекламации [5, С.6]. Актуальна разработка системы поддержки принятия решений на основе правил доказательной аргументации, чтобы повысить эффективность обработки заявок.

Результаты исследования

Запрос предназначен для выбора информации из базы/хранилища данных/знаний и ее последующей обработки по заданному правилу/правилам.

Рассмотрим схему обработки запросов в СППР на рисунке 1.



Рисунок 1 – Типовая схема обработки запроса

На рисунке 1 приняты следующие обозначения цифрами:

- 1 – формирование текста запроса;
- 2 – формирование параметров запроса для реализации условия отбора при классификации;
- 3 – запуск на выполнение запроса;
- 4 – получение выборки по результатам выполнения запроса;
- 5 – выгрузка результата выполнения запроса;
- 6 – построение сводной таблицы по результатам выполнения запроса.

Для классификации запросов предлагается классификатор, на основе которого строится схема взаимодействия интерфейса пользователя с сервером базы данных (рисунок 2).

Схематично процедуры извлечения и выбора признаков с использованием технологии автоматического машинного обучения представлены на рисунке 3.

У запроса есть числовые и текстовые характеристики [9, С.268]. К числовым характеристикам относятся следующие параметры: интенсивность запросов, время пересылки и обработки запросов, объёмы запросов и ответов на запросы. Методики измерения и оптимизация этих параметров подробно описаны в [1, С.46;6, С.55].

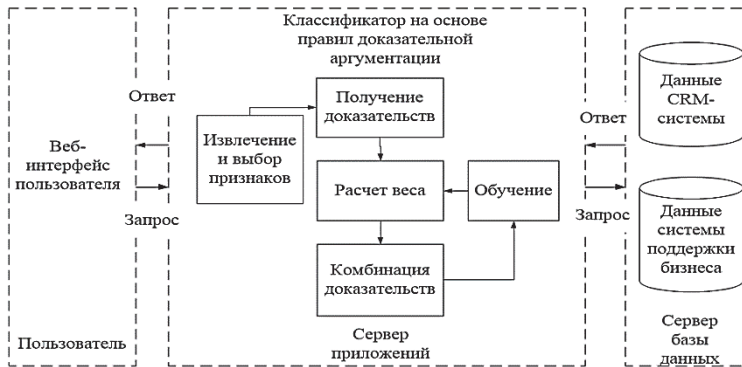


Рисунок 2 – Схема взаимодействия интерфейса пользователя с сервером базы данных с помощью классификатора запросов

На входе классификатора, представленного на рисунке 2, характеристики запросы (числовые и текстовые). На выходе классификатора – прогнозируемые степени уверенности (обычно обозначаются p_{ni}), соответствующие заранее определенному классу запроса.

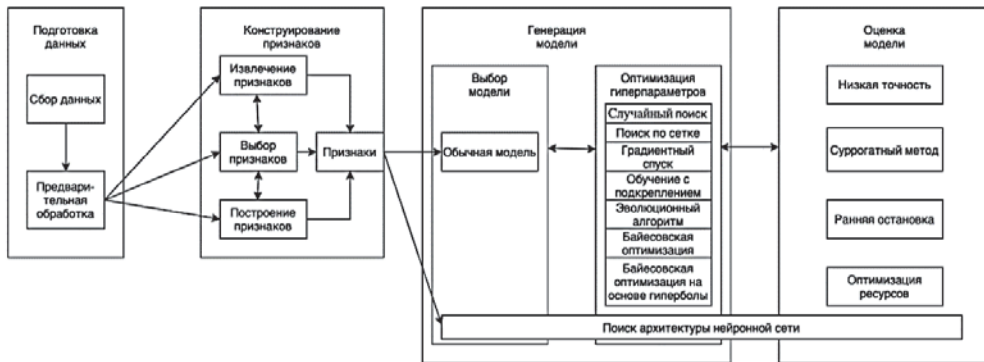


Рисунок 3 – Схема конструирования признаков и генерации модели оптимизации параметров запросов

На рисунке 3 показаны пути оптимизации параметров запросов с помощью классификатора на основе правил доказательной аргументации для повышения производительности обработки.

Для того чтобы охарактеризовать запрос, сначала выбираются текстовые признаки. Отобразить текстовые характеристики запроса в математическом виде можно с помощью математической модели векторного пространства, показанной в геометрической интерпретации на рисунке 4.

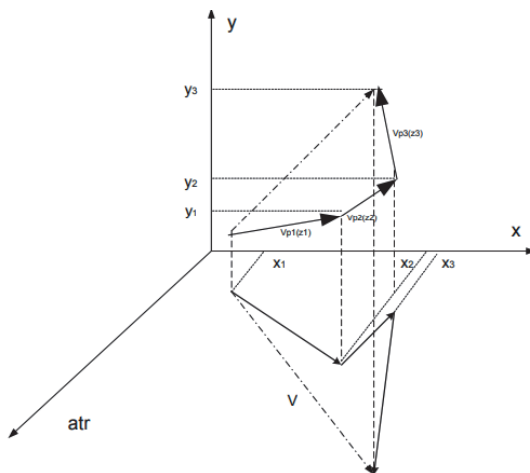


Рисунок 4 – Геометрический смысл векторного представления текста запроса

Модель, представленная на рисунке 4, трехмерная, характеризует координаты x_i и y_i вектора цели; atr – атрибут, который имеет каждая координата. Единицы измерения координат вектора: слова, именные группы, отдельные предложения, смысловые группы предложений, абзацы и т.д. Атрибуты координаты x_1 : $atr_x =$ (во все времена), атрибуты координаты y_1 : $atr_y =$ (одни и те же).

Для определения весов векторов (то есть количества компонент, которые равны 1) применяется метод TF-IDF (от англ. TF – term frequency, IDF – inverse document frequency – статистическая мера, используемая для оценки важности слова в контексте документа) [8, С.488]. Оценка по методу TF-IDF находится в диапазоне значений [0;1].

В правилах доказательной аргументации надежность и вес доказательств определяются отдельно. В подходе доказательной аргументации для многокритериальных задач принятия решений оценка альтернативы по каждому критерию рассматривается как свидетельство, а вес свидетельства равен весу критерия. В результаты классификации вносят больший вклад те доказательства (свидетельства), которые являются более важными [10, С.181].

Доказательство (свидетельство) зависит от значений связанных свойств, его можно представить следующим образом

$$e_\gamma = \{(y_n, p_N), n=1, 2, \dots, N\}, \quad (1)$$

где y_n – класс запроса; p_N ($p_N = 1/N$) – степень доверия (уверенности) в соответствии с определенным признаком; N – количество классов запросов.

Пусть $d_{j\gamma}$ обозначает евклидово расстояние между свидетельством e_j и равномерным распределением e_γ .

$$d_{j\gamma} = d(e_j, e_\gamma) = \sqrt{\sum_{n=1}^N (p_{n,j} - p_N)^2}, j = 1, \dots, L, d_{j\gamma} \in [0,1], \quad (2)$$

где L – количество ссылочных значений A_{ij} ($j=1, \dots, L$).

Из (2) можно сделать вывод, что при $N \rightarrow \infty$ $d_{j\gamma} \rightarrow 0$. Чем больше евклидово расстояние, тем выше важность доказательства e_j . Таким образом, важность доказательства e_j увеличивается, при увеличении евклидова расстояния $d_{j\gamma}$.

Чтобы оценить стоимость признаков запроса, разделяют номинальную стоимость и числовую. Пусть для характеристики запроса x используются M признаков. Если эти признаки номинальные, то каждая характеристика x_i ($i=1, \dots, M$) имеет L ссылочных значений A_{ij} ($j=1, \dots, L$), и можно выделить N возможных классов, которым может принадлежать запрос. Соответственно, имеется $M \cdot L$ доказательств. Определенный образец запроса пользователя x активирует M доказательств e_i ($i=1, \dots, M$).

При определении числовой стоимости для определенного образца запроса пользователя $x = (x_1, \dots, x_i, \dots, x_M)$ с числовыми функциями, функция x_i активирует два соседних элемента доказательства e_j^i с активированным весом $\alpha_{i,j}$ и e_{j+1}^i с активированным весом $\alpha_{i,j+1}$, если признак x_i принимает свое значение в интервале $[A_{ij}^i, A_{j+1}^i]$. Преобразование информации используется для генерации активированных весов [7, С.28].

$$\alpha_{i,j} = \frac{A_{j+1}^i - x_i}{A_{j+1}^i - A_j^i} \text{ и } \alpha_{i,j+1} = 1 - \alpha_{i,j}, \text{ при } A_j^i \leq x_i \leq A_{j+1}^i, i = 1, \dots, M, j = 1, \dots, L, \quad (3)$$

$$\alpha_{i,j'} = 0, \quad j' = 1, \dots, L, \quad j' \neq j, j + 1.$$

Из (3) следует

$$e_i = \{(y_n, p_{n,i}), n = 1, \dots, N, i = 1, \dots, M\}, \quad (4)$$

$$p_{n,i} = \alpha_{i,j} p_{n,j} + \alpha_{i,j+1} p_{n,j+1}, j = 1, \dots, L$$

В (4) p_{ni} – это степень уверенности, с которой класс считается u_n , учитывая, что функция x_i активирует e_j^i и e_{j+1}^i одновременно [11, С.183]. Между тем, вес w_i и надежность r_i доказательства e_i могут быть вычислены как взвешенная сумма таковых двух соседних частей доказательства. То есть

$$w_i = \alpha_{i,j} w_j^i + \alpha_{i,j+1} w_{j+1}^i \text{ и } r_i = \alpha_{i,j} r_j^i + \alpha_{i,j+1} r_{j+1}^i, i = 1, \dots, M, j = 1, \dots, L. \quad (5)$$

Правила доказательной аргументации используют для объединения активированных свидетельств с их с весами соотношение

$$O(x) = \{(y_n, p_{n,e(M)}), n=1, \dots, N\}, \quad (6)$$

где определенная выборка x принадлежит определенному классу u_n , кото-

рый имеет максимальную степень доверия $p_{n,e(M)}$.

Практическое применение исследований: проектирование системы поддержки принятия решений на основе правил доказательной аргументации для оператора сотовой связи. Проведен анализ случаев, когда абоненты опрошенного оператора мобильной связи обращались с жалобами на услуги связи. Различные заявки на определенные услуги классифицированы и переданы для решения в отдел технической поддержки.

Типичный процесс рассмотрения заявок в мобильной связи заключается в следующем. Когда администраторы в центре обслуживания клиентов получают заявку клиента, они записывают ее и решают, является ли заявка действительной, исходя из своего опыта или использования CRM-систем. Если они смогут обработать заявку и разрешить ее самостоятельно, они отправят отзыв непосредственно клиенту. В противном случае они передадут заявку в соответствующие отделы сервисной поддержки. Технические специалисты отделов проанализируют причины жалоб, предложат решения и отправят отзывы клиентам. При вынесении решения о причине жалоб технические специалисты должны эффективно использовать как информацию из систем поддержки бизнеса, так и описания жалоб клиентов. Затем заявки будут разделены на две категории, включая класс качества сети (NQC) и класс абонентского терминала (CTC) в зависимости от причины жалоб.

Образцы данных включают подробную информацию о запросах пользователей, результатах обработки и соответствующих отзывах клиентов. В запросах исследуемой компании выделено шесть числовых характеристик:

- (1) D – поступает ли заявка из области, где действуют помехи сигнала;
- (2) E – поступает ли заявка из места скопления людей;
- (3) CT – поступает ли заявка в период скопления людей;
- (4) BS – исправное ли рабочее состояние текущей базовой станции;
- (5) B – поступает ли заявка из области со слабым сигналом;
- (6) MP – существует ли проблема совместимости между сетью мобильной связи и терминалом, используемым подающим заявку пользователем.

Все функции представлены тремя ссылочными значениями {ноль, один, неизвестно/ unknown}, где ноль означает «НЕТ», один – «ДА», а unknown представляет отсутствующую или неизвестную информацию.

Из-за того, что полученная информация часто противоречива, техническим специалистам сложно делать точные суждения, что приводит к неточности классификации. Сбор данных включает описания заявок, которые направлены клиентами в отдел технической поддержки телекоммуникационной компании, период – три месяца. Всего после устранения некоторых недействительных заявок получено 1433 записи в качестве выборочных данных. Образцы данных содержат подробную информацию о заявках пользователей, результатах обработки и соответствующих отзывах клиентов. Типичная запись о заявке показывает время и место сбоя обслуживания, описание клиентом восприятия сетевых услуг, таких как медленный доступ в Интернет и/или низкое качество связи.

Поскольку значений шести ключевых характеристик, по которым прово-

дился анализ, может быть недостаточно для технических специалистов для принятия точного решения, текстовые элементы извлекаются из описаний заявки, чтобы охарактеризовать заявку, а затем отбираются для классификации вместе с шестью ключевыми характеристиками, указанными выше. Метод TF-IDF используется для извлечения текстовых функций. В соответствии со значениями коэффициента передачи информации для функций при классификации выбираются десять основных текстовых функций, как показано в таблице 1. Из таблицы 1 видно, что значения TF-IDF этих десяти текстовых функций для каждого описания заявки являются непрерывными, они будут дискретизированы до сбора доказательств. Их ссылочные значения установлены равными $\{0, 0,33, 0,66\}$ в соответствии с интервалом данных.

Доказательства для каждой особенности получены с помощью байесовской статистики. Возьмем, к примеру, функцию D . D указывает, поступила ли заявка из района, где действуют помехи сигнала. Поскольку значение признака D дискретно, его ссылочное значение $A_j^i = \{1, 0, \text{unknown}\}$.

Таблица 1 – Основные текстовые функции

ID функции	Текст	Коэффициент передачи информации
F1	Вечер	0.0108
F2	Веб-страница	0.0104
F3	Функция	0.0103
F4	Открыть	0.0102
F5	Только	0.0101
F6	Начинать	0.0100
F7	Отказ	0.0098
F8	Скорость	0.0098
F9	Данные	0.0097
F10	Вызов	0.0010

Три части свидетельства, относящиеся к признаку D , могут быть описаны следующим образом:

$$e_1^D = \{(NQC, 1), (CTC, 0)\}, \tag{7}$$

$$e_2^D = \{(NQC, 0.4916), (CTC, 0.5084)\}, \tag{8}$$

$$e_3^D = \{(NQC, 0), (CTC, 0)\}. \tag{9}$$

Вероятность c_{nj} и степень достоверности p_{nj} признака D могут быть получены с использованием уравнений (7) и (8), как показано в таблице 2. Это означает, что если значение признака D заявки равно 1, вероятность того, что заявка принадлежит NQC , составляет 100%. В противном случае заявка принадлежит NQC с вероятностью 49,16% и CTC с вероятностью 50,84%.

Если информация об объектах является неполной из-за отсутствия данных в записях, расчет вероятности будет незнанием. В схеме правил доказательной аргументации незнание представлено степенью уверенности, присвоенной атрибуту со значением «неизвестно».

Таблица 2 – Вероятность и степень уверенности признака D

		Референтные значения D		
		1 (e_1)	0 (e_2)	Неизвестный (e_3)
Правдоподобие	NQC	0,033	0,967	0
	CTC	0	1	0
Степень уверенности	MTNQC	1	0,4916	0
	CTC	0	0,5084	0

Возьмем, к примеру, функцию MP . Его ссылочное значение A_j^i является одним из значений в наборе $\{1,0, \text{unknown}\}$, и свидетельство, связанное с каждым из ссылочных значений признака, выглядит следующим образом.

$$e_{1}^{MP} = \{(NQC, 0.6826), (CTC, 0.3174)\},$$

$$e_{2}^{MP} = \{(NQC, 0.2439), (CTC, 0.7561)\},$$

$$e_{3}^{MP} = \{(NQC, 0.0971), (CTC, 0.9029)\}.$$

Для подтверждения веса и надежности определяется распределение $e_\gamma = \{(y_n, p_N)\}$, $n=1, 2, \dots, 16\}$, которое представляет собой свидетельство без какой-либо возможности классификации. Поскольку всего для классификации используются 10 текстовых и 6 числовых признаков, p_N является константой со значением $1/16$. Вес каждого доказательства определяется уравнением (3). Доказательства, связанные с шестью числовыми характеристиками и их весами, перечислены в таблице 3.

Таблица 3 – Характеристики доказательств

Функция	Значение	Свидетельство	NQC	CTC	Вес
D	1	e_1^D	1.0000	0.0000	0.8496
	0	e_2^D	0.4900	0.5100	0.4717
	unknown	e_3^D	0.0000	0.0000	0.2355
F	1	e_1^F	0.2360	0.7640	0.6012
	0	e_2^F	0.6309	0.3692	0.5063
	unknown	e_3^F	0.0000	0.0000	0.2355
CT	1	e_1^{CT}	0.0000	1.0000	0.8498
	0	e_2^{CT}	0.6794	0.3205	0.5354
	unknown	e_3^{CT}	0.0000	0.0000	0.2358
BS	1	e_1^{BS}	0.4912	0.5087	0.4717
	0	e_2^{BS}	0.9004	0.0998	0.7368
	unknown	e_3^{BS}	0.0000	0.0000	0.2358
B	1	e_1^B	0.9600	0.0400	0.8035
	0	e_2^B	0.3295	0.6706	0.5297
	unknown	e_3^B	0.0000	0.0000	0.2357
MP	1	e_1^{MP}	0.6825	0.3172	0.5374

Функция	Значение	Свидетельство	NQC	СТС	Вес
	0	e_2^{MP}	0.2438	0.7562	0.5943
	unknown	e_3^{MP}	0.09712	0.9028	0.7394

Надежность свидетельства изначально устанавливается равной его весу в контексте классификации, а затем обучается на основе выборочных данных. Функция *Fmincon* в наборе инструментов оптимизации Matlab's используется для поиска оптимальной надежности каждой части свидетельства. Эта функция минимизирует евклидово расстояние между реальным выходом и объединенным выходом классификатора на основе правил доказательной аргументации.

После того, как запрос/заявка пользователя поступает в систему поддержки принятия решений, основанную на доказательствах, каждое свидетельство, относящееся к функциям, активируется и объединяется рекурсивной аналитической моделью правил доказательной аргументации. В исследовании используется матрица неточностей для измерения эффективности классификатора на основе правил доказательной аргументации. Матрица неточностей представляет собой таблицу с двумя строками и двумя столбцами, которые представляют количество истинно положительных (*TP*), ложноположительных (*FP*), ложноотрицательных (*FN*) и истинно отрицательных (*TN*) классов заявки, как показано в таблице 4.

Таблица 4 – Матрица неточностей

		Прогнозируемый класс	
		NQC	СТС
Актуальный класс заявки	NQC	Истинно положительный (TP)	Ложноотрицательный (FN)
	СТС	Ложноположительный (FP)	Истинно отрицательный (TN)

Если запрос/заявка, которая фактически принадлежит NQC, прогнозируется как NQC, она засчитывается как истинно положительная; если она классифицируется как СТС, он считается ложноотрицательным. Если запрос/заявка, фактически принадлежащая СТС, классифицируется как NQC, она считается ложноположительной; если он классифицируется как СТС, он считается истинно отрицательным. Введены четыре показателя эффективности: точность A_c , отзывчивость (или чувствительность) S_e , прецизионность P_r и *F*-мера F_m , которые рассчитываются в эксперименте по обработке M запросов посредством десятикратной перекрестной проверки.

$$A_c = \frac{TP+TN}{M}, \tag{10}$$

$$S_e = \frac{TP}{TP+FN}, \tag{11}$$

$$P_r = \frac{TP}{TP+FP}, \tag{12}$$

$$F_m = \frac{2}{\frac{1}{Pr} + \frac{1}{Se}}. \quad (13)$$

Точность A_c – это измерение того, насколько часто классификатор верен. Напоминание, известное как чувствительность S_e или показатель истинных положительных результатов, предназначено для измерения того, как часто классификатор предсказывал NQC , когда заявка пользователя фактически принадлежит NQC . Прецизионность P_r , также называемая положительной прогностической ценностью, – это количество NQC , деленное на общее количество элементов, помеченных как принадлежащие NQC . F-мера F_1 (или оценка F1) – это мера точности исследования с учетом как прецизионности, так и отзыва. TP – истинно положительные классы запросов. TN – истинно отрицательные классы запросов.

Для эксперимента выбраны десять текстовых и шесть числовых функций. Десятикратная перекрестная проверка проводится для оценки надежности предложенной модели классификации. Модель обучалась и тестировалась соответственно десять раз. Средние значения десятикратной перекрестной проверки (т.е. десять значений) по четырем параметрам производительности, описанным выше, показаны в таблице 5.

Таблица 5 – Средние значения результатов десятикратной перекрестной проверки

	Точность	Прецизионность	Отзыв	F-мера
Обучение	0.8591	0.9259	0.8416	0.8817
Тестирование	0.8527	0.9188	0.8383	0.8762

Эффективность классификации предлагаемого классификатора на основе правил доказательной аргументации может варьироваться в зависимости от количества функций. Чтобы повысить предсказательную способность модели классификации, ниже анализируются характеристики модели с различными текстовыми и числовыми функциями. Сравнение показателей эффективности показано на рисунке 5. Буквой D по оси ординат на рисунке 5 обозначена вероятность и степень уверенности признака.

Количество текстовых функций изменено с десяти до пятидесяти для сравнения. Эксперименты показывают, что классификация шестнадцати функций дает в целом высокую производительность. Это также указывает на то, что необходимо извлекать текстовые элементы из описаний клиентов для лиц, принимающих решения, хотя в описаниях заявок присутствует неопределенная и противоречивая информация.

Для построения классификаторов часто используются классические инструменты вероятностного моделирования, включая логистическую регрессию (LR), байесовские сети (BN) и наивный байесовский алгоритм (NB), а также невероятностные методы, включая опорную векторную машину (SVM) и дерево решений $J48$. Кроме того, для расчета принимаются типичные классификаторы ансамбля, включая случайный лес, упаковку в мешки и бустинг (adaboost).

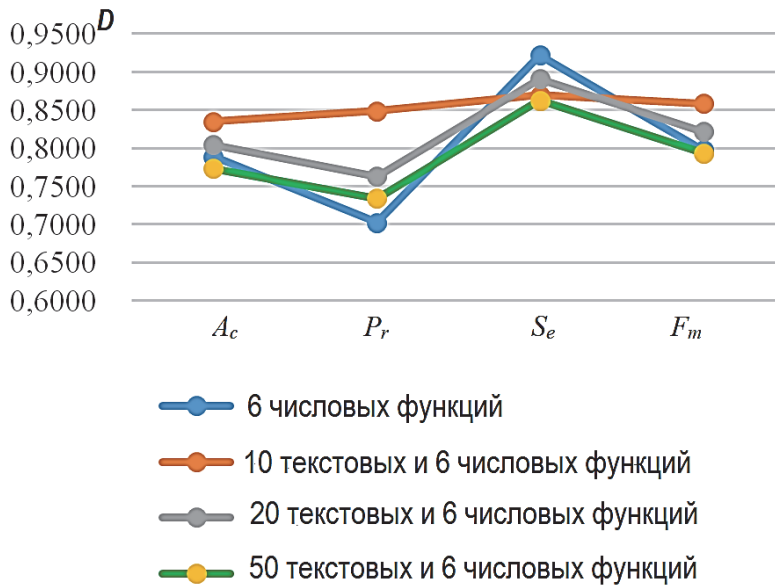


Рисунок 5 – Производительность классификации с разным количеством признаков

Чтобы сравнить эффективность их классификации с производительностью предложенного классификатора на основе правил доказательной аргументации, все девять классификаторов обучаются с использованием одного и того же набора данных. За исключением классификатора на основе правил доказательной аргументации, реализованного в программном обеспечении Matlab, остальные семь классификаторов разрабатываются в программном обеспечении WEKA, инструментарии интеллектуального анализа данных с открытым исходным кодом. Все параметры алгоритмов обучения восьми классификаторов принимают значения по умолчанию в программном обеспечении WEKA.

Эффективность восьми классификаторов с 16 признаками показана в таблице 6. Показатели эффективности указывают на то, что классификатор на основе правил доказательной аргументации имеет конкурентоспособные характеристики по сравнению с другими классификаторами, когда для классификации используются как текстовые, так и числовые признаки. Имеется небольшое улучшение значений точности классификатора на основе правил доказательной аргументации, по сравнению с вероятностным классификатором и с ансамблевым классификатором.

Из сравнительного анализа видно, что предлагаемая СППР на основе доказательной аргументации имеет нижеприведенные преимущества.

Во-первых, она обеспечивает управляемый данными и вероятностный инструмент моделирования, когда все доказательства, их вес и надежность сформулированы на основе реальных данных.

Таблица 6 – Эффективность девяти классификаторов с шестнадцатью характеристиками

	Точность	Прецизионность	Отзыв	F-мера
Классификация на основе правил доказательной аргументации	0.8559	0.9221	0.8400	0.8789
Сеть Байеса	0.8570	0.8570	0.8600	0.8550
Наивный байесовский классификатор	0.8530	0.8530	0.8570	0.8520
Логистическая регрессия	0.8420	0.8420	0.8420	0.8420
Метод опорных векторов	0.8000	0.8000	0.8440	0.7870
Дерево решений (J48)	0.8591	0.8590	0.8751	0.8550
Бустинг (Adaboost)	0.8300	0.8300	0.8310	0.8280
Случайный лес	0.8540	0.8540	0.8751	0.8490

Во-вторых, он может агрегировать противоречивые доказательства в условиях неопределенности для классификации. Свидетельство определяется распределением убеждений по силовому набору структуры различения. Естественная и гибкая форма распределения вероятностей позволяет делать неточные рассуждения на любом уровне абстракции.

В-третьих, его можно использовать для объединения разнородных данных из нескольких источников для принятия решений.

Одним из возможных ограничений является то, что как структурированные данные из систем поддержки принятия решений, так и неструктурированные данные из описаний заявок/запросов должны быть преобразованы в доказательства для принятия решений.

Выводы

Разработана система поддержки принятия решений, основанная на доказательной аргументации, для обработки заявок пользователей в мобильной связи. Неопределенная информация формулируется как распределение убеждений, а роли различных характеристик различаются посредством соответствующих свидетельств с разной степенью надежности и веса.

Предлагаемый классификатор доказательной аргументации, основанный на правилах, снабжен внутренними функциями для обработки недостающих данных без удаления или вменения.

В эмпирическом исследовании выбираются шесть числовых и десять текстовых характеристик, относящихся к качеству связи, для характеристики заявки потребителя мобильной связи. Описан систематический процесс рассмотрения

заявок и выделены входы-выходы предлагаемого классификатора.

Экспериментальные результаты показывают его высокую точность классификации, которая конкурирует с классическими вероятностными моделями. Производительность предлагаемого классификатора на основе правил доказательной аргументации превосходит некоторые классификаторы ансамбля, такие как классификаторы посредством «Мешков» и бустинг (Adaboost).

Предлагаемая интеллектуальная СППР на базе Интернета предоставляет техническим специалистам по телекоммуникациям информативную и основанную на знаниях методологию систематической и автоматической обработки заявок клиентов.

Литература

1. Аббасова Т.С., Артюшенко В.М., Акимкина Э.Э. Повышение эффективности систем поддержки принятия решений на основе многомерных хранилищ данных: монография. Москва-Берлин. Издательство: ООО «Директ-Медиа». 2021. 128 с.
2. Аббасова Т.С., Ганюшина Ю.Н., Ганюшин А.А. Автоматизация делопроизводства организации / Матрица научного познания. 2022. № 5-1. С. 14-27.
3. Аббасова Т.С., Гурин Д.В., Ребрунова А.И., Аббасов Э.М. Построение эластичной промышленной среды на собственных аппаратных ресурсах для облачных приложений // Эволюционные процессы информационных технологий. Сборник научных статей по материалам 8-й Международной научно-технической конференции 4 апреля 2023. Бургас: Изд-во ИГНЭИТ, 2023. С. 4-12 (178 с.).
4. Аббасова Т.С., Златов М.М., Чибисова О.В., Петровская З.В. Исследование характеристик автоматизированного рабочего места сетевого программиста // Эволюционные процессы информационных технологий. Сборник научных статей по материалам 8-й Международной научно-технической конференции 4 апреля 2023. Бургас: Изд-во ИГНЭИТ, 2023. С.13-19 (178 с.).
5. Аббасова Т.С., Чибисова О.В. Разработка модуля клиент-серверного приложения для оптимизации работы оперативно-диспетчерской группы // Информационно-технологический вестник. 2022. № 1(31). С. 3-13.
6. Артюшенко В.М., Акимкина Э.Э. Имитационная модель адаптивной системы поддержки принятия решений // Вестник компьютерных и информационных технологий. 2018. № 2. С. 46-56.
7. Бурлаева Е.И., Павлыш В.Н. Проект построения алгоритма классификации текстовых документов // Проблемы искусственного интеллекта. 2017. № 4(7). С. 24-31.
8. Егоров Д.П., Егоров А.И., Федотов А.А., Кравченко О.В. Классификация текстовых сообщений с помощью TF-IDF и word2vec // Цивилизация знаний: российские реалии. Сборник трудов XXI Международной научной конференции. 2020. С. 487-490.

9. Кривошеев Н.А., Спицын В.Г. Методы машинного обучения для классификации текстовой информации // Труды Международной конференции по компьютерной графике и зрению «Графикон». 2019. № 29. С. 266-269.
10. Оськина К.А. Оптимизация метода классификации текстов, основанного на TF-IDF, за счет введения дополнительных коэффициентов // Вестник Московского государственного лингвистического университета. 2016. № 15(754). С. 175-187.
11. Янина А.О., Воронцов К.В. Мультимодальные тематические модели для разведочного поиска в коллективном блоге // Машинное обучение и анализ данных. 2016. Т. 2. № 2. С. 173-186.