УДК 51-73+ 004.852

Елпатова Александра Юрьевна, студент,
Амурский государственный университет

ПРОГРАММНЫЙ МОДУЛЬ ФОРМИРОВАНИЯ ВЕРОЯТНОСТЕЙ ДЛЯ БАЙЕСОВСКОЙ СЕТИ ДИАГНОСТИКИ ИНФЕКЦИОННЫХ ЗАБОЛЕВАНИЙ ЛЁГКИХ

Аннотация. В статье рассматривается архитектура и алгоритмическая основа программного модуля формирования вероятностей, являющегося вычислительным ядром системы поддержки принятия врачебных решений при диагностике инфекционных заболеваний лёгких. Описаны компоненты модуля – загрузчик таблиц условных вероятностей (CPTLoader) и движок вероятностного вывода (InferenceEngine), реализующие точный вывод.

Ключевые слова: Байесовская сеть доверия, вероятностный вывод, маргинализация, таблицы условных вероятностей, поддержка принятия врачебных решений, ФЗ-152.

Разработка интеллектуальных систем поддержки принятия врачебных решений является одним из приоритетных направлений цифровой трансформации здравоохранения. В отличие от детерминированных экспертных систем, байесовские сети доверия позволяют формализовать вероятностную природу медицинских знаний и выполнять рассуждения в условиях неопределённости, неполноты клинических данных и противоречивости симптомов [1]. Настоящая работа посвящена описанию программного модуля формирования вероятностей, реализованного в рамках системы диагностики инфекционных заболеваний лёгких (пневмония, бронхит, туберкулёз и др.).

Модуль спроектирован в соответствии со стандартами IEEE 830-1998 и ГОСТ 19.201-78 и реализован на языке Python 3.10+ с использованием библиотеки pgmpy. Архитектура модуля построена по принципу разделения ответственности и включает пять ключевых компонентов (таблица 1).

Таблица 1

Внутренние компоненты модуля

Компонент	**Назначение**
ProbabilityModule (API)	Точки входа, внешний программный интерфейс
CPTLoader	Чтение таблиц условных вероятностей из JSON/YAML/CSV и априорных вероятностей из шифрованной БД
NetworkBuilder	Формирование направленного ациклического графа (DAG) и TabularCPD
InferenceEngine	Вероятностный вывод (Variable Elimination), маргинализация
Validator	Проверка нормировки CPT, ацикличности графа, допустимости свидетельств
HypothesisRanker	Ранжирование диагнозов и выделение ключевых предикторов

Процесс инициализации начинается с состояния «Не инициализирован». Компонент CPTLoader обращается к файловой системе для чтения конфигурационных файлов (форматы JSON, YAML, CSV), содержащих таблицы условных вероятностей, и к локальной шифрованной базе данных SQLite+SQLCipher (AES-256) для получения априорных

вероятностей диагнозов. Собранные данные передаются компоненту NetworkBuilder, который формирует направленный ациклический граф (DAG) и объекты TabularCPD. Компонент Validator выполняет обязательную проверку нормировки (сумма вероятностей по каждому родительскому состоянию равна 1,0 ± ε) и ацикличности графа. При успешной валидации модуль переходит в состояние «Готов к работе».

Основной рабочий сценарий – расчёт апостериорных вероятностей. Он инициируется событием infer (evidence). Основное приложение передаёт словарь клинических свидетельств вида {node_id: state}, включающий категориальные (наличие/отсутствие симптома), числовые (значения лабораторных показателей, температура) и текстовые данные. Ключевой особенностью модуля является встроенная поддержка неполных данных: если часть клинических показателей отсутствует, компонент InferenceEngine автоматически выполняет маргинализацию – суммирование совместного распределения по ненаблюдаемым узлам сети до запуска вывода. Это соответствует требованию технического задания об обработке неполных данных и позволяет системе функционировать в реальных клинических условиях, когда не все обследования доступны врачу на момент принятия решения.

Точный вероятностный вывод выполняется методом Variable Elimination через библиотеку pgmpy. Алгоритм последовательно исключает ненаблюдаемые переменные из совместного распределения, что обеспечивает вычислительную эффективность даже для сетей средней размерности (до 50 узлов – время вывода не превышает 1 секунды). Полученные апостериорные вероятности P (D|e) нормализуются (приведение к диапазону [0; 1] с суммой, равной 1,0) и передаются компоненту HypothesisRanker, который сортирует диагнозы по убыванию вероятности и выполняет анализ чувствительности для выделения ключевых предикторов – переменных, оказавших наибольшее влияние на итоговый результат.

Важнейшим аспектом проектирования модуля является обеспечение соответствия требованиям Федерального закона № 152-ФЗ «О персональных данных». Модуль функционирует исключительно в локальном контуре, не осуществляет сетевых обращений и не хранит персональные данные пациентов – он получает только обезличенный набор клинических свидетельств. База знаний с априорными вероятностями хранится в шифрованной локальной БД (SQLCipher, AES-256), что исключает несанкционированный доступ к медицинским правилам.

Архитектура модуля предусматривает возможность актуализации базы медицинских знаний без перекомпиляции исходного кода. Сценарий update_cpt (new_cpt) позволяет врачу-эксперту или методисту редактировать таблицы условных вероятностей через внешние конфигурационные файлы. Модуль автоматически выполняет валидацию нормировки, при успехе сохраняет обновлённые CPT и инициирует перестроение байесовской сети в памяти без перезапуска процесса.

Таким образом, разработанный программный модуль формирования вероятностей обеспечивает математически обоснованный вероятностный вывод в байесовских сетях доверия с поддержкой неполных клинических данных. Модульная архитектура, локальное шифрование базы знаний и возможность обновления медицинских правил без перекомпиляции делают систему перспективным инструментом для внедрения в клиническую практику лечебных учреждений с соблюдением требований российского законодательства о персональных данных.

Список литературы:

1. Рассел С., Норвиг П. Искусственный интеллект: современный подход / пер. с англ. – Москва: Вильямс, 2016. – 1408 с.

2. ГОСТ 19.201-78. Единая система программной документации. Техническое задание. Требования к содержанию и оформлению. – М.: ИПК Издательство стандартов, 2001.

3. IEEE 830-1998 Recommended Practice for Software Requirements Specifications. – IEEE, 1998.

4. Федеральный закон от 27.07.2006 № 152-ФЗ «О персональных данных» (с изм. и доп.).

5. Документация библиотеки pgmpy: Python library for Probabilistic Graphical Models [Электронный ресурс]. – URL: https://pgmpy.org (дата обращения: 12.11.2026).