

Задаючись такими критеріями, як швидкість та простота розробки, а також наявність існуючих рішень для стандартних процедур, для програмної реалізації методики класифікації емоційної забарвленості тексту було обрано бібліотеки – NumPy, Tensorflow, Keras, Matplotlib, Sklearn, мову програмування Python.

Варто зазначити, що розроблена методика для вирішення задачі класифікації емоційної забарвленості тексту відрізняється легкістю реалізації, високою точністю і швидкістю, але в той же час потребує невелику кількість ресурсів для навчання. Крім того, отримані результати точності моделі, не є остаточними. Дана модель може бути дотренована на нових даних задля збільшення точності класифікації або адаптації моделі для іншої предметної області.

Висновки. В роботі було створено методику для класифікації емоційної забарвленості тексту на основі методу машинного навчання з вчителем. На етапі проектування було виконано аналіз існуючих методів автоматичного аналізу емоційної забарвленості тексту та обґрунтування моделі вирішення поставленої задачі. Для реалізації було використано мову програмування Python, технології машинного навчання та нейронних мереж. Тестування методики класифікації емоційної забарвленості тексту розробленої на базі рекурентної нейронної моделі LSTM на даних з датасету IMDB підтвердило її працездатність.

Список використаної літератури

1. “Natural Language Processing. What it is and why it matters”, стаття. [Online]. Available: https://www.sas.com/en_us/insights/analytics/what-is-natural-language-processing-nlp. [Accessed: September 09, 2021].
2. “Sentiment Analysis Based on Deep Learning: A Comparative Study”, стаття. [Online]. Available: <https://www.mdpi.com/2079-9292/9/3/483/pdf>. [Accessed: September 09, 2021].
3. “IMDB Dataset”, сторінка датасету. [Online]. Available: <http://ai.stanford.edu/~amaas/data/sentiment/>. [Accessed: September 09, 2021].
4. “Glove word embeddings”, стаття. [Online]. Available: <https://nlp.stanford.edu/projects/glove/>. [Accessed: September 09, 2021].

УДК 004.891.3

БАЗОВЫЕ ПОНЯТИЯ ИНТЕЛЛЕКТУАЛЬНЫХ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ ОБРАБОТКИ СИГНАЛОВ

Файнзильберг Л.С. (fainzilberg@gmail.com)

*Международный научно-учебный центр информационных технологий и систем
НАН Украины и МОН Украины (Україна)*

Дано определение термина «интеллектуальная информационная технология обработки сигналов». Рассмотрен подход к построению таких технологий с помощью инструментальной системы СИДИГРАФ.

В последние годы в научной и научно-популярной литературе все чаще упоминаются термины «информационные технологии» (ИТ) и «интеллектуальные ИТ. И это не случайно, поскольку и те и другие оказывают существенную роль в решении актуальных прикладных задач современного общества, в частности в цифровой медицине [1].

В то же время, необоснованное применение этих терминов к прикладным системам, которые не являются ни ИТ, ни тем более интеллектуальными ИТ, может привести к дискредитации этих важных научных направлений, как это уже не раз было в прошлом, например, к дискредитации термина «АСУ – автоматизированные системы управления».

Поскольку наука начитается с определений, попытаемся дать более-менее строгие формулировки понятиям «ИТ обработки сигналов» и «интеллектуальные ИТ обработки сигналов», оттолкнувшись от базового определения самого термина «технология», которое *объединяет* два понятия: «техно» и «логия» (рис. 1).



Рис. 1. Общее понятие термина «ТЕХНОЛОГИЯ»

Таким образом, *в общем случае*, технология представляет собой искусство преобразования некоторого сырья в продукт (например, железной руды в сталь), а наука технологии направлена на поиск *наиболее эффективных* способов такого преобразования.

Отсюда можно сформулировать такие определения:

Определение 1. ИТ обработки сигналов – совокупность вычислительных процедур, обеспечивающих эффективный переход от наблюдаемого сигнала, который порождает объект исследования (сырье), к информации о текущем состоянии объекта (продукт).

Определение 2. Интеллектуальная ИТ обработки сигналов – совокупность вычислительных процедур, обладающих *свойствами естественного интеллекта*, которые обеспечивают эффективный переход от наблюдаемого сигнала к информации о текущем состоянии объекта.

Размышления об естественном интеллекте приводят к выводу о том, что интеллектуальная ИТ должна обладать, *по крайней мере*, такими свойствами:

- *обобщение* – способность классифицировать наблюдаемые сигналы;
- *обучаемость* – способность улучшать свои характеристики по мере эксплуатации;
- *адаптация* – способность приспосабливаться к неизвестным и изменяющимся ситуациям внешней среды;
- *инвариантность* – нечувствительность к возмущениям внешней среды;
- *прогнозирование* – возможность предсказывать будущие ситуации по наблюдению только частей (фрагментов) наблюдаемого сигнала;
- *понимание* – способность интерпретировать сигнал на основе сравнения его текущих характеристик с прошлыми значениями;
- *гибкость* – устойчивость к возможным неудачам и способность коррекции принимаемых решений;
- *взаимозаменяемость* – использование альтернативных методов анализа сигналов;
- *коммуникабельность* – способность предоставлять информацию в форме, понятной для восприятия конкретным пользователем с учетом его квалификации.

Возможны формальный и неформальный подходы к построению эффективных процедур, реализующих отдельные этапы ИТ, в том числе, интеллектуальных процедур. Первый подход основан на решении оптимизационной задачи: вычислительная процедура строится *формально* из условия минимума (максимума) некоторого критерия (суперкритерия) \mathfrak{Z} , характеризующего эффективность конкретного этапа обработки.

Например, байесовская схема классификации основана на минимуме критерия \mathfrak{Z} , представляющего собой средний риск принимаемых решений, а процедура построения

регрессионной модели – на основе минимума критерия \mathfrak{Z} , представляющего собой среднеквадратическую ошибку аппроксимации экспериментальных данных функцией, заданной с точностью до неизвестных параметров.

Однако не все вычислительные процедуры могут быть построены на основе формального подхода. Далеко не всегда удается выразить критерий \mathfrak{Z} эффективности конкретного этапа обработки сигнала в виде выпуклой функции, для которой можно провести глобальную оптимизацию. Можно указать и другие причины ограничений формального подхода. В таких ситуациях не остается ничего иного, как строить вычислительную процедуру неформально на основе интуиции, а критерий \mathfrak{Z} использовать лишь для оценки ее эффективности.

Пусть \mathfrak{Z}_0 – априорное значение критерия \mathfrak{Z} отдельного этапа ИТ, которое может быть оценено (формально или экспертом) до использования некоей вычислительной процедуры \mathfrak{R} , а \mathfrak{Z}_1 – апостериорное значение критерия \mathfrak{Z} , оцененное после использования этой процедуры.

Определение 3. Процедура \mathfrak{R} эффективна, если выполняется *строгое* неравенство:

$$\mathfrak{Z}_1 < \mathfrak{Z}_0. \quad (1)$$

Определение 4. Процедура $\mathfrak{R}^{(i)}$ более эффективна, чем процедура $\mathfrak{R}^{(j)}$, если для апостериорных значений соответствующих критериев справедливо *строгое* неравенство:

$$\mathfrak{Z}_1^{(i)} < \mathfrak{Z}_1^{(j)}. \quad (2)$$

Таким образом, при неформальном построении эффективной процедуры не обязательно решать оптимизационную задачу поиска глобального минимума. Более того, в соответствии с (1), (2) процедура остается эффективной, даже если она не удовлетворяет *локальному* минимуму критерия \mathfrak{Z} , а лишь позволяет уменьшить \mathfrak{Z} по сравнению с его априорным значением. Разумеется сформулированные определения естественным образом могут быть переформулированы, если процедура \mathfrak{R} направлена на увеличение значения \mathfrak{Z} .

Для практической реализации предложенного подхода разработана инструментальная система СИДИГРАФ (СИгнал-ДИалог-ГРАФик), которая обеспечивает интерактивный синтез интеллектуальных ИТ обработки сигнала [2]. Система содержит открытый для расширения набор вычислительных компонент, реализующих отдельные стадии ИТ. Выбор и настройка компонент осуществляет конструктор ИТ в режиме диалога с системой.

Для придания системе определенной гибкости ее ядро построено на двух программных классах – обобщенной модели обработки данных (МОД) и обобщенной модели носителя данных (МНД).

Обобщенная модель обработки данных представляет собой *абстрактный класс* с универсальными полями, свойствами и методами, *инвариантными* относительно конкретной процедуры обработки. При этом МОД содержит виртуальный метод «пустой» обработки, который доопределяется в наследниках этого класса.

Обобщенная модель носителя данных представляет собой *абстрактный класс*, на базе которого создаются объекты (наследники), обеспечивающие хранение цифровых данных и доступ к ним через порождаемые графические образы (рис. 2).

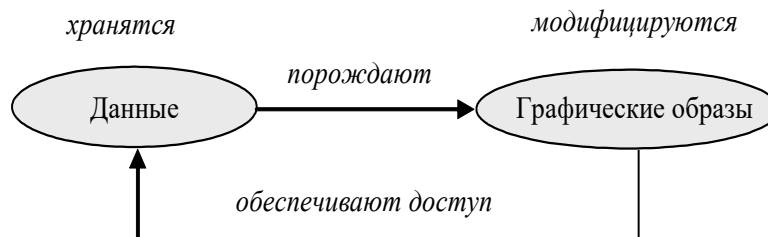


Рис. 2 Концептуальная идея построения класса МНД

Инструментальная система позволила разработать ряд интеллектуальных ИТ для решения задач в разных сферах, в том числе, разработать комплекс ФАЗАГРАФ, реализующий инновационный в кардиологии метод обработки электрокардиограмм [3]. Опыт разработки этих технологий показал, что в рамках *алгоритмического подхода* построение интеллектуальных ИТ предполагает *активное* участие разработчика технологии, который, используя свой естественный интеллект, создает эффективные вычислительные процедуры (рис. 3).

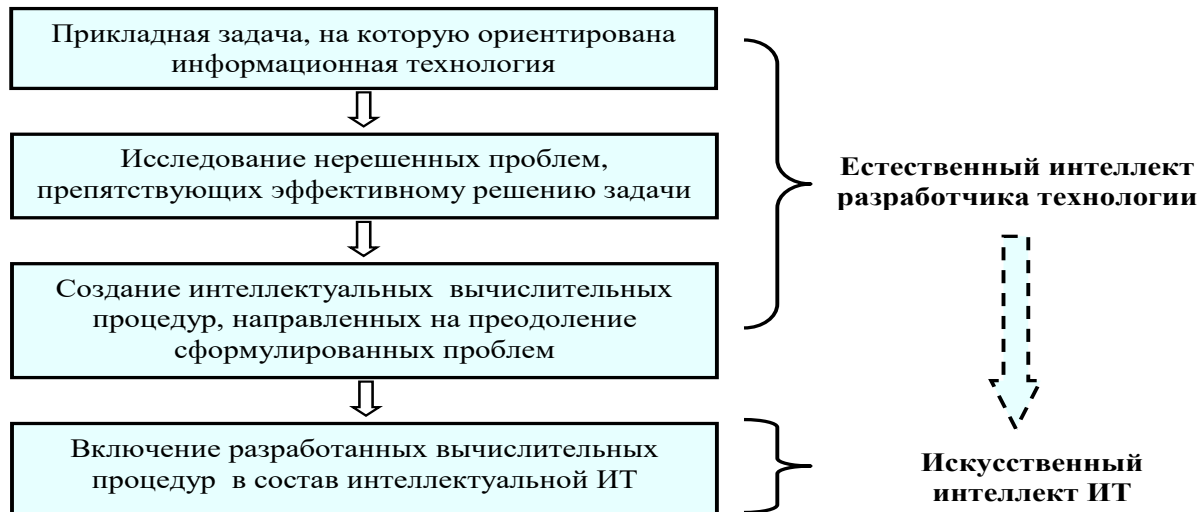


Рис. 3. Этапы построения интеллектуальной ИТ

Список использованной литературы

1. Гриценко В.И., Файнзильберг Л.С. Интеллектуальные информационные технологии в цифровой медицине на примере фазаграфии. Киев: Наукова Думка, 2019. 423 с.
2. Файнзильберг Л.С. Интерактивный синтез информационных технологий обработки сигналов с локализованной информацией // Кибернетика и вычислительная техника. 2017. № 1 (187). С. 11-29.
3. Файнзильберг Л.С. Основы фазаграфии. Киев: Освіта України, 2017. 264 с.