

Волчкова Екатерина Сергеевна, магистрант
ФГАОУ ВО «МГТУ «СТАНКИН»

АЛГОРИТМ ОПРЕДЕЛЕНИЯ МЕТАПРОГРАММ СОИСКАТЕЛЯ ВАКАНТНОЙ ДОЛЖНОСТИ НА ОСНОВЕ АНАЛИЗА ТРАНСКРИБАЦИИ ИНТЕРВЬЮ

Аннотация. В статье предложен алгоритм автоматического определения метапрограмм кандидата на основе транскрибации интервью. Алгоритм включает извлечение языковых индикаторов, их взвешивание и расчет метапрограммных осей. Дополнительно рассчитываются индекс достоверности и показатель соответствия кандидата требованиям вакансии.

Ключевые слова: Метапрограммы, анализ интервью, обработка естественного языка, рекрутинг.

В современных системах управления персоналом возрастает роль автоматизированных методов анализа информации о кандидатах. Одним из источников таких данных являются интервью, содержащие сведения о профессиональном опыте, мотивации и особенностях мышления кандидата. Для описания когнитивных стратегий человека в HR-практике используются метапрограммы – устойчивые способы восприятия информации и принятия решений [1].

Традиционно определение метапрограмм выполняется специалистом на основе анализа ответов кандидата. Однако такой подход является трудоемким и субъективным. Развитие технологий обработки естественного языка позволяет применять алгоритмические методы анализа транскрибаций интервью и выявлять поведенческие паттерны автоматически [2].

В данной работе предлагается алгоритм определения метапрограмм кандидата на основе анализа транскрибации интервью. Исходными данными могут являться аудио или видеозаписи интервью, которые на первом этапе преобразуются в текст с помощью систем автоматического распознавания речи. Полученная транскрибация передается языковой модели для извлечения индикаторов метапрограмм.

Индикаторы представляют собой текстовые фрагменты высказываний кандидата, которые могут свидетельствовать о проявлении определенных когнитивных стратегий. Каждый индикатор характеризуется двумя параметрами: силой проявления и степенью уверенности модели. Сила индикатора отражает степень выраженности признака и задается дискретным значением от 1 до 3. Уверенность модели принимает значение в диапазоне от 0 до 1.

Для количественной оценки индикаторов используется взвешенный показатель, который рассчитывается по формуле (1):

$$\text{indicator_score} = \text{strength} * \text{confidence}, \quad (1)$$

где strength – силы выраженности признака, а confidence – уверенность модели.

Использование произведения этих параметров позволяет учитывать одновременно как выраженность признака, так и надежность его определения моделью. Подобный подход применяется в задачах интеллектуального анализа данных при агрегировании признаков различной достоверности [3].

Для повышения устойчивости анализа вводится фильтрация индикаторов по минимальному пороговому значению. Если значение indicator_score меньше 0,5, соответствующий индикатор исключается из дальнейшего анализа, что позволяет уменьшить влияние слабых или случайных признаков.

После фильтрации выполняется расчет силы каждой метапрограммы с использованием модели накопления доказательств, где суммируются веса всех относящихся к ней индикаторов. Итоговый расчет производится по формуле (2):

$$\text{meta_score} = \sum \text{indicator_score}, \quad (2)$$

где indicator_score – рассчитанный ранее взвешенный показатель.



Такой подход позволяет учитывать не только количество обнаруженных индикаторов, но и степень их выраженности. В результате формируется количественная оценка каждой метапрограммы.

Для интерпретации результатов метапрограммы объединяются в бинарные оси, отражающие противоположные когнитивные стратегии. Например, одной из таких осей является «процесс – результат». Ориентация на процесс характеризуется вниманием к последовательности действий и выполнению этапов работы, тогда как ориентация на результат связана с фокусом на достижении конечной цели.

Для определения положения кандидата на оси используется нормализованная функция (3):
$$\text{axis_index} = (\text{right_score} - \text{left_score}) / (\text{right_score} + \text{left_score} + 1), \quad (3)$$

где right_score и left_score представляют собой суммарные значения метапрограмм, относящихся к соответствующим сторонам оси. Подобная нормализация используется в задачах сравнительного анализа признаков, поскольку позволяет учитывать относительное соотношение показателей [4].

Добавление единицы в знаменатель выполняет роль сглаживающего коэффициента. Это предотвращает нестабильность результатов в случаях, когда количество индикаторов невелико. Значение показателя axis_index находится в диапазоне от -1 до $+1$.

Для удобства интерпретации результатов значение оси переводится в дискретную шкалу по формуле (4):

$$\text{axis_value} = \text{round}(\text{axis_index} \times 2), \quad (4)$$

где axis_index – рассчитанное ранее положение кандидата на оси метапрограммы.

В результате получаются значения от -2 до $+2$, где отрицательные значения указывают на преобладание одной стороны оси, положительные – противоположной, а значение 0 отражает относительный баланс характеристик.

Дополнительным элементом алгоритма является расчет индекса достоверности полученного профиля. Индекс учитывает количество обнаруженных индикаторов и среднюю уверенность модели. Расчет выполняется по формуле (5):

$$R = \text{avg_confidence} \times \ln(N + 1), \quad (5)$$

где avg_confidence – средняя уверенность индикаторов, а N – их количество. Использование логарифмической функции позволяет уменьшить влияние очень большого числа индикаторов на итоговый показатель. В зависимости от значения индекса можно выделить низкую, среднюю и высокую степень достоверности результатов анализа.

Алгоритм также предусматривает возможность оценки соответствия кандидата требованиям вакансии. Для этого метапрограммный профиль кандидата сравнивается с целевым профилем позиции. Расхождение между значениями метапрограммных осей преобразуется в показатель соответствия, который агрегируется с учетом важности каждой характеристики.

Предложенный алгоритм реализован в виде программного модуля и может использоваться в системах интеллектуального анализа интервью кандидатов. Использование подобных методов позволяет повысить объективность оценки кандидатов и сократить время анализа информации в процессе подбора персонала

Список литературы:

1. О'Коннор Д., Сеймор Дж. Введение в нейролингвистическое программирование. – 2-е изд. – Челябинск: Библиотека А. Миллера, 1998. – 266 с.
2. Speech and Language Processing. – 3rd ed. draft. изд. – [Электронный ресурс]. – Режим доступа: <https://web.stanford.edu/~jurafsky/slp3/>
3. Jiawei Han, Micheline Kamber, Jian Pei Data Mining: Concepts and Techniques. – 3rd ed. изд. – Waltham, MA: Morgan Kaufmann Publishers, 2012. – 702 с.
4. Рассел С., Норвиг П. Искусственный интеллект: современный подход. – 2-е изд. – Москва: Издательский дом “Вильямс”, 2006. – 1409 с

