

Навигатор профессий и навыков

В условиях быстро меняющегося рынка труда людям все сложнее ориентироваться в выборе профессии и оценивать свои карьерные возможности. Кроме того, с развитием технологий и автоматизации процессов многие профессии уходят в прошлое, а новые возникают, требуя от соискателей новых навыков и знаний, и понимания того, как их человеческий потенциал может быть эффективно реализован. Создание Навигатора является важным шагом в процессе повышения прозрачности рынка труда для всех заинтересованных агентов – работников, работодателей, государственных органов. В современном быстроменяющемся мире под воздействием цифровизации запрос на новые навыки регулярно обновляется. При этом понимание профессий остается более инертным. Навигатор позволит создать ориентиры для работников на разных этапах их трудовой жизни по выстраиванию индивидуальных образовательных и карьерных траекторий и достижения успеха в меняющемся мире.

Навигатор профессий и навыков – это автоматизированный инструмент, обученный на обширном массиве данных рынка труда с использованием методов обработки естественного языка, включая большие языковые модели (LLM), который поможет пользователям оценить релевантность своих навыков набору профессий и предоставит рекомендации тому, какие навыки необходимо освоить для изменения карьерной траектории.

Навигатор профессий и навыков может также использоваться в образовательных целях для актуализации содержания образовательных программ. Он поможет в позиционировании и маркетинге образования, что поможет повысить качество отбора и мотивацию обучающихся.

Основой Навигатора является обширная база данных, включающая более 35 миллионов вакансий российского рынка труда за период с 2021 по 2023 годы. В процессе анализа данные были обработаны с применением методов обработки естественного языка (NLP), таких как RuBERT и Sentence BERT, что позволило объединить различные написания профессий и выделить 1522 уникальные профессии в 32 профессиональных областях.

Для задачи точного выделения навыков из текстов вакансий была дообучена языковая модель Mistral 7B, включающая 7 миллиардов параметров. Модель прошла обучение на базе данных, состоящей из 25 тысяч вакансий, и применялась для инференса на выборке из 707 тысяч вакансий. Процесс инференса осуществлялся с использованием удалённых серверов и параллелился на нескольких графических процессорах (GPU).

Различия в написании навыков в текстах вакансий и их синонимичность потребовали разработки методологии по их агрегированию. Предложенная методология применялась к

объединенным по профессиональным сферам вакансиям и выделенным в них навыкам в два этапа.

На первом этапе проводилось автоматизированное объединение навыков с помощью поиска наиболее близких вариантов их написания по расстоянию Дамерау-Левенштейна. В качестве «основного написания» навыка использовался наиболее часто встречающийся вариант, а остальные варианты заменялись им. Затем проводилась экспертная оценка автоматизированного объединения.

На втором этапе была проведена экспертная работа по объединению наиболее часто встречаемых навыков по их смыслу для каждой профессиональной сферы

Затем для расчета зарплатных премий навыков был применен двухэтапный алгоритм. На первом этапе, с помощью модели коррекции Хекмана, определяется вероятность того, что в объявлении указана заработная плата. На втором этапе, с учетом определенной вероятности и используя метод гедонического ценообразования и модель LASSO, оценивается, как наличие отдельных навыков влияет на размер заработной платы, что и называется зарплатной премией. При этом чем выше зарплатная премия, тем выше ценность навыка для работодателей.

Таким образом, Навигатор профессий и навыков представляет собой мощный инструмент для получения актуальной информации о профессиях и построения карьерных траекторий, основываясь на анализе большого объема данных о рынке труда.

Разработка и тестирование модели проводились с применением суперкомпьютерного комплекса НИУ ВШЭ, а проект реализовывался научными подразделениями НИУ ВШЭ: Лабораторией исследований рынка труда и Международной лабораторией экономики нематериальных активов.