

ЗАКЛЮЧЕНИЕ
временной экспертной комиссии по рассмотрению документов
образовательной программы «Машинное обучение в цифровом продукте»
(направление подготовки 01.04.02 Прикладная математика и информатика;
уровень высшего образования: магистратура)

Сформированная УМС (протокол № 041 (061) от 25.12.2024) временная экспертная комиссия (ВЭК) по рассмотрению поступивших в УМС НИУ ВШЭ документов образовательной программы «Машинное обучение в цифровом продукте» (направление подготовки 01.04.02 Прикладная математика и информатика; магистратура) в составе *Дегтярева К.Ю.* (руководитель комиссии), *Кольцова С.Н.* (член комиссии), *Архимандритова И.Б.* (член комиссии) и *Ясницкого Л.Н.* (член комиссии) изучила представленный разработчиками пакет документов ОП.

В целом, члены ВЭК согласились с тем, что программа представляется актуальной, соответствует приоритетам развития НИУ ВШЭ и имеет все шансы быть востребованной на рынке образовательных услуг.

Одновременно, члены ВЭК представили свои комментарии и сформулировали замечания, которые можно свести к следующему:

1. В перечень проанализированных разработчиками магистерских программ (документ «Обоснование открытия образовательной программы») не попала значимая программа «Машинное обучение и анализ данных», которая реализуется в Санкт-Петербургском кампусе ВШЭ. В списке программ-конкурентов достаточно слабо отражены преимущества заявленной программы перед программами-конкурентами – по сути, всё сравнение сводится к перечислению того, чем характеризуются соотв. программы (их описание);

2. Как было отмечено членами ВЭК, в пул обязательных дисциплин входит дисциплина «Теория вероятностей и математическая статистика», но в этом же пуле отсутствует, например, дисциплина «Линейная алгебра», которая не менее важна в контексте данной программы. Возможно, стоило бы предусмотреть несколько математических дисциплин по выбору для тех студентов, которым необходимо (в частности, входное тестирование могло бы это подтвердить) «освежить» или структурировать знания в соответствующей области перед изучением дисциплин специализации;

3. В аннотациях дисциплин отсутствуют указания пререквизитов, которые крайне важны для понимания «выстраиваемых» цепочек изучения материала, наполнения дисциплин материалом, основанным на ранее изучаемых темах;

4. В описании дисциплины «Программирование на Python» рекомендуется указать более «свежую» литературу – здесь можно опираться на новые и уже достаточно известные книги, например, «Объектно-ориентированный Python» (авторы С.Ф. Лотт, Д. Филлипс), изд-во 'Питер', 2024 г., «Python Data Structures Pocket Primer» (автор О. Campesato), Mercury Learning and Information, 2024 г., «Python 3 and Feature Engineering» (автор О. Campesato), 2024 г. (при этом, имеется целый ряд переводных книг по языку Python 2022-2024 гг. издания, которые представляют углубленные аспекты программирования).

Кроме этого, в описании дисциплины «Обработка естественного языка» указана литература двухлетней давности – нельзя сказать, что за 2 года соотв. источники потеряли свою актуальность, но за этот период в данной области произошли существенные изменения и, как следствие, можно отметить появившиеся совсем недавно издания, которые могли бы быть полезны при подготовке соотв. дисциплины – например, «LLM Engineer’s Handbook» (авторы P. Iusztin, M. Labonne), Packt Publishing Ltd., 2024 г., «Introduction to Python and Large Language Models. A Guide to Language Models» (автор D. Grigorov), APress, 2024 г., «Transformers for Natural Language Processing and Computer Vision» (автор D.Rothman), 3-е изд., Packt Publishing Ltd., 2024 г. Фактически, большие языковые модели изучаются в рамках трех дисциплин программы – складывается некоторое начальное непонимание, как соотносятся соответствующие темы охватываемого материала в разных дисциплинах;

5. Дисциплина «Введение в MLOps», по сути, представляет собой введение в DevOps с некоторыми дополнениями, касающимися собственно MLOps (3 пункта содержания дисциплины из 12-ти имеющихся) и основных принципов построения web-сервисов (здесь наблюдается пересечение с дисциплинами по backend-разработке и «продвинутому» Python). Возможно, стоило бы исключить из содержания дисциплины темы по web-разработке в пользу более глубокого изучения аспектов MLOps, жизненного цикла систем машинного обучения (напр., распределенные системы хранения, noSQL, Data Lake и Data Warehouse, масштабируемость и сокращение времени обучения за счёт применения параллельного и распределенного обучения, и т.п.). В дисциплину, связанную с backend-разработкой, стоило бы включить изучение одного из самых быстрых и популярных фреймворков FastAPI, который занимает лидирующие позиции на рынке;

6. Было также отмечено, что в рамках дисциплины «GPU и распределенные вычисления» стоило бы изучить темы, связанные с основными фреймворками для инференса (процесса работы, получения конечного результата обработки данных) на GPU (TensorRT SDK, ONNX Runtime, и др.);

7. По мнению членов ВЭК, в дисциплине «Продвинутый Python» основное внимание уделяется не столько средствам языка, сколько практикам DevOps, что перекрывается с материалом дисциплины «Введение в MLOps». Возможно, основное внимание стоило бы уделить изучению именно продвинутых возможностей современных версий языка, как это следует из названия дисциплины, а также использованию инструментальных средств, необходимых в сфере машинного обучения;

8. Было высказано мнение, что дисциплина «Алгоритмы на Python» примерно соответствует классической дисциплине по алгоритмам для студентов 1-2 курсов бакалавриата; в данном же случае, речь идёт о дисциплине магистерской программы, поэтому имело бы смысл больше внимания уделить специфическим алгоритмам, применяемым в области машинного обучения;

9. Обсуждался и вопрос, связанный с целесообразностью указания тех фреймворков, которые планируются к использованию в рамках некоторых дисциплин. Например, в дисциплине «Глубинное обучение» это может быть PyTorch, часто используемая «связка» высокоуровневого API Keras с платформой Tensorflow, или набирающий популярность фреймворк JAX. То же относится и к дисциплине «Машинное обучение» (пакет scikit-learn, активно использующийся в сходных учебных дисциплинах?); по мнению членов ВЭК, в описании содержания последней из перечисленных дисциплин также отсутствует важный раздел, связанный с оптимизацией комбинации гиперпараметров классических алгоритмов. Если в дисциплинах «Глубинное обучение в обработке звука и видео» и «Компьютерное

зрение» планируется использование готовых трансформеров с Hugging Face, то рекомендуется это указать в соотв. описаниях дисциплин.

Руководитель и члены комиссии представили заполненные чек-листы с соответствующими оценками вынесенного на рассмотрение пакета документов ОП по различным параметрам со своими замечаниями и комментариями. После подсчета выставленных членами ВЭК оценок, средний балл по всем критериям оказался равен **63,9**.

В результате проведенного обсуждения члены ВЭК большинством голосов приняли решение согласовать представленные документы ОП. Временная экспертная комиссия рекомендует УМС поддержать программу «Машинное обучение в цифровом продукте» (направление подготовки: 01.04.02 Прикладная математика и информатика; уровень высшего образования: магистратура) и рекомендовать ее к рассмотрению на УС НИУ ВШЭ.

При этом, члены ВЭК рекомендуют разработчикам обратить внимание на представленные комментарии и замечания (пп. 1-9) на последующих этапах работы с документами ОП, внесения в них изменений.