**О т з ы в**

на магистерскую диссертацию

**«Разработка моделей предупреждения сбоев на производственной линии**

**с использованием методов машинного обучения»**

**Прудникова Евгения Александровича**

В настоящее время активно используются системы, построенные на основе алгоритмов машинного обучения. В магистерской диссертации поставлена задача исследования методов машинного обучения в применении к задачам обработки данных с последующей разработкой и реализацией модели на основе методов машинного обучения для предотвращения сбоев на производственной линии. Тема диссертации актуальна в силу того, что производственные линии снабжаются большим числом датчиков, позволяющими собирать большой объем информации о производственном процессе, а последующая обработка этой информации и разработка на ее основе встраиваемой модели, несомненно, будут способствовать предупреждению сбоев на производственной линии, что в перспективе приведет к повышению эффективности производства.

В первой главе диссертации выполнен краткий обзор физических и программных систем предупреждения ошибок на производственных линиях. Во второй главе приведено подробное описание методов машинного обучения (supervised learning, unsupervised learning, semi-supervised learning, transductive learning, reinforcement learning, online learning, active learning, meta-learning/ learning-to-learn) и сделан вывод о том, что для решения поставленных в рамках магистерской диссертации задач, наиболее подходящей методикой является использование именно методов машинного обучения (обучение с учителем)*.* В третьей главе описаны средства разработки: *Apache Spark* (программный каркас с открытым исходным кодом для реализации распределённой обработки неструктурированных и слабоструктурированных данных, входящий в экосистему проектов *Hadoop)* иязык программирования *Python.* Так как для построения хорошей модели необходимо провести детальный анализ данных (зависимости, корреляции, добавление новых признаков для улучшения качества, удаление признаков, вносящих шум), то это и было выполнено автором магистерской диссертации в главе четыре: отобраны признаки, имеющие наименьшую корреляцию между собой, а также наивысшую корреляцию с целевой переменной. В пятой главе Прудников Е.А. обосновывает выбор алгоритма машинного обучения (градиентный бустинг (англ. *boosting* — улучшение) над деревьями) для его использования в построенной модели. Для построения модели Прудниковым Е.А. использовалась готовая реализация алгоритма в свободной библиотеке xgboost, а для сравнения качества алгоритма был взят алгоритм машинного обучения – *RandomForest* (реализация взята из *Python*-библиотеки для машинного обучения *sklearn)*. Результаты работы алгоритмов (оценка качества алгоритма и метрика качества MCC) в магистерской диссертации представлены в таблицах 5.1 и 5.2 (полученное качество модели по метрикам *МСС* и *AUC-PRC* – 0.89) и на основе анализа полученных результатов в работе был сделан вывод о том, что классификатор *Random Forest* показывает лучшие результаты.

В процессе работы над магистерской диссертацией Прудников Е. А. показал высокий уровень владения современными техническими инструментами. Новизна магистерской диссертации состоит в том, что интегрирование построенной в работе модели в систему производственной линии позволит своевременно обрабатывать измерения датчиков производственной линии и, исходя из этого, предсказывать возникновение сбоев, что позволит заранее принять меры к их устранению и, как следствие, снизить издержки производства и удешевить продукт для конечного пользователя. Важной особенностью модели является то, что она учитывает совокупность показаний датчиков, не привязываясь к понятию норма для конкретного датчика, также в построенная модель основана на современных методах машинного обучения, которые доказали свою эффективность в настоящее время. К недостаткам указанной работы можно отнести отсутствие публикаций (выступлений на конференции, актов внедрения) Прудникова Е.А. Магистерская диссертация соответствует требованиям, предъявляемым к магистерским диссертациям, материал изложен последовательно и ясно.

Магистерская диссертация прошла проверку системой Антиплагиат: заимствований 15.57%, цитирований - 0%, оригинальность – 84.43%.

Считаю, что магистерская диссертация «Разработка моделей предупреждения сбоев на производственной линии с использованием методов машинного обучения» Прудникова Евгения Александровича заслуживает оценки **8 (восемь),** а ее автору может быть присвоена степень «Магистр математики и информационных технологий».

Руководитель,

кандидат физ.-мат. наук,

доцент

12 июня 2017 г. Соболевская Е.П.