

ПОДДЕРЖКА ПРИНЯТИЯ РЕШЕНИЙ В УПРАВЛЕНИИ ПРОИЗВОДСТВОМ НА ОСНОВЕ ИНТЕЛЛЕКТУАЛЬНОЙ ОБРАБОТКИ ДОКУМЕНТАЦИИ И ЧИСЛОВЫХ ДАННЫХ

Петрова Т. А.¹, Шоев Р. Н.¹

Научный руководитель – доктор техн. наук, доцент Духанов А. В.¹

¹Университет ИТМО

petrova.tatiana.itmo@mail.ru

Работа выполнена в рамках темы НИР «Разработка метода интеллектуального моделирования производственного процесса на основе анализа документации с применением больших языковых моделей».

Введение

Современные производственные системы характеризуются высокой вариативностью технологических параметров, ростом требований к качеству продукции, а также чувствительностью качества целевого продукта к технологическим параметрам производства: температурным режимам, составу используемого сырья, настройкам оборудования, времени выдержки и другим факторам. Большая часть знаний о процессе каждого конкретного производства преимущественно зафиксирована в неструктурированном формате в виде технологических регламентов, операционных карт, отчетов, протоколов испытаний. Ручной анализ такого рода данных требует существенных временных затрат, а конечный результат в значительной степени зависит от экспертного опыта инженеров. В то же время проведение производственных испытаний для поиска оптимальных рецептов чаще всего оказывается ресурсозатратным, что является препятствием для проведения новых экспериментов [1].

В последние годы большие языковые модели (БЯМ) демонстрируют значительный потенциал в задачах обработки специализированных технических текстов. Обзор применения БЯМ в промышленной индустрии показывает эффективность их использования для анализа производственной документации и в системах поддержки принятия инженерных решений [2]. Российские исследовательские центры активно развивают методы автоматизации анализа технической документации. Так, например, цифровой корпоративный эксперт «Акела», разработанный учеными Университета ИТМО совместно с промышленными партнерами, демонстрирует способность дообучать большие языковые модели на предметно-ориентированных данных предприятия и консультировать специалистов в диалоговом режиме по профессиональным вопросам [3]. В зарубежной научной практике также существуют примеры применения БЯМ для обработки технической информации. К примеру, в области химической инженерии разрабатываются системы на основе БЯМ для интерактивного взаимодействия с пользователями и ускорения технологических исследований [4].

Однако большинство существующих решений ориентированы либо на диалоговые консультации по общим вопросам, либо на автоматическое извлечение заданных фактов из текстовых документов [5], без последующей интеграции с моделями машинного обучения и непосредственного применения к задачам оптимизации производства. Таким образом, остается актуальной задача создания целостного интеллектуального приложения, которое способно не только выполнить анализ документации, но и определить производственные ограничения, смоделировать технологический процесс и оптимизировать его параметры под конкретные инженерные требования.

Основная часть

В качестве решения предлагается прототип интеллектуального приложения, обеспечивающего переход от анализа текстового запроса пользователя к оперативному формированию рекомендаций по оптимизации параметров производства. Основная идея данного подхода заключается в использовании гибридной архитектуры, объединяющей три ключевых модуля: NLP-модуль на основе дообученной большой языковой модели для автоматического структурирования текстовых технических заданий, предсказательную модель для прогнозирования свойств материалов, модуль оптимизации для подбора оптимальных технологических параметров с учетом производственных ограничений.

Реализованный программный прототип представлен в виде локального приложения и апробирован на данных отечественного предприятия по производству многослойных керамических конденсаторов на примере задачи оптимизации формирования керамических пленок. Выбор разработки локального приложения обусловлен требованиями компании к безопасности корпоративных данных и защиты их от утечек.

На вход в NLP-модуль поступает текстовый запрос с техническим заданием от технолога. Разработаны две версии приложения, которые могут быть использованы в зависимости от имеющихся ресурсов и потребностей. Упрощенная версия использует дообученную на размеченных данных русскоязычную модель ruBERT для извлечения заданных именованных сущностей из запросов пользователей [6]. Система распознает в запросах технологов ключевые технологические ограничения (такие как номер партии материала, температура, влажность и т.д.), преобразует их в формат, пригодный для модели прогнозирования, и передает в модуль оптимизации. Расширенная версия приложения основана на легковесной модели Qwen и способна обрабатывать более сложные и объемные текстовые запросы, обучаться на приложенных пользователем технических документах, предоставлять консультации в формате общения с моделью в чате, отвечая на уточняющие вопросы технолога, что значительно расширяет сферу применения системы.

Помимо поля для ввода запроса в интерфейсе предусмотрено управление настройками модуля оптимизации: выбор как одного, так и нескольких целевых параметров для максимизации итоговых значений, выбор алгоритма оптимизации и количество итераций, оказывающих влияние на время ожидания ответа. В рамках текущей реализации доступны два метода: байесовская оптимизация (ТРЕ) для быстрого получения ответа в течение нескольких секунд, и генетический алгоритм для более глубокого глобального поиска оптимума. Выбор пользователем стратегии позволяет учитывать компромисс между скоростью расчета и точностью оптимизации, что является важным фактором в промышленной среде с ограничениями по времени принятия решений.

За определение суррогатной модели для задач однокритериальной и многокритериальной оптимизации отвечает модуль прогнозирования, внутри которого осуществляется подбор наиболее подходящей под соответствующую задачу регрессионной модели с подобранными по сетке гиперпараметрами.

Приложение протестировано на исторических данных предприятия по производству многослойных керамических конденсаторов. В настоящее время ведутся работы по проведению опытной эксплуатации разработки в производственном контуре компании.

Выводы

В данной работе представлена разработка метода интеллектуального моделирования производственного процесса на основе анализа технической документации с применением больших языковых моделей, машинного обучения и

алгоритмов оптимизации. Практическая значимость предложенного решения заключается в сокращении времени анализа документации, снижении риска получения дефектной продукции и обеспечении выбора оптимальных инженерных решений. Проведение опытной эксплуатации позволит количественно оценить экономический эффект и определить направления дальнейшего масштабирования метода.

Литература

1. Pei X., Pei J., Hou H., et al. Optimizing casting process using a combination of small data machine learning and phase-field simulations // *npj Computational Materials*. 2025. Vol. 11. Article 27.
2. Добренко Н.В., Добренко Д.А., Улизько М.В. Интеллектуальная поддержка принятия управленческих решений в MES-системах с использованием больших языковых моделей // *Экономика. Право. Инновации*. 2024. № 3. С. 45-52.
3. Цифровой корпоративный эксперт «Акела» [Электронный ресурс] // ТАСС. 26.09.2023. URL: <https://tass.ru/tehnologii/18845699> (дата обращения: 20.02.2026).
4. Zhang H., Zhou J., Xu F., et al. A large language model system for the field of chemical engineering technology // *arXiv*. 2025. 2509.07034. URL: <https://arxiv.org/abs/2509.07034> (дата обращения: 20.02.2026).
5. Nabavi A., Ramaji I., Sadeghi N., Anderson A. Leveraging Natural Language Processing for Automated Information Inquiry from Building Information Models // *Journal of Information Technology in Construction (ITcon)*. 2023. Vol. 28. P. 266–285.
6. Zmitrovich D., Abramov A., Kalmykov A., et al. A Family of Pretrained Transformer Language Models for Russian // *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*. Torino, Italy, 2024. P. 507–524.