Кейс 2 «НЛМК»

Кейс содержит в себе описание внедрения в работу Новолипецкого металлургического комбината (НЛМК) технологий искусственного интеллекта (ИИ) с разных точек зрения: 1) НЛМК и 2) группы ИТ-разработчиков.

Работа с кейсом проводится в два этапа:

Этап 1:

- изучение материалов кейса (Приложение 1),

- ответ на четыре проверочных вопроса (Таблица 1). Таблица 1. Задание на письменную работу студентам

|  |  |  |  |
| --- | --- | --- | --- |
| **№** | **Вопрос** | **Ответ** | **Аргументация** |
| 1 | Укажите направление в разработке искусственного интеллекта, который применялся в рассматриваемом кейсе. | Направления ИИ, выберите правильный вариант ответа:   1. Игры и машинное творчество 2. Интеллектуальные роботы 3. Системы обработки естественного языка и чат боты 4. Нейронауки и распознавание образов 5. Системы, основанные на   знаниях и экспертные системы | Приведите аргументы в пользу выбранного ответа |
| 2 | Укажите какие элементы входят в систему ИИ в НЛМК. |  | Приведите аргументацию своего ответа (не  более 500 знаков) |
| 3 | В деятельности каких сотрудников НЛМК, помимо сталеваров может применяться  разработанный ИИ? |  | Приведите аргументацию своего ответа (не  более 500 знаков) |
| 4 | Может ли НЛМК оптимизировать свои затраты на персонал путем сокращения сталеваров и  использования ИИ? |  | Приведите аргументацию своего ответа (не  более 500 знаков) |

Этап 2.

Найдите как развивается НЛМК после начального этапа внедрения систем искусственного интеллекта. Какие технологии еще проходили и проходят внедрение? С какими сложностями сталкиваются? Как решаются?

Ответ представить в виде презентации.

# Новостное сообщение

**НЛМК внедрил систему искусственного интеллекта в помощь сталевару** Источник: [https://lipetsk.nlmk.com/ru/media-center/press-releases/nlmk-vnedril-sistemu-](https://lipetsk.nlmk.com/ru/media-center/press-releases/nlmk-vnedril-sistemu-iskusstvennogo-intellekta-v-pomoshch-stalevaru/) [iskusstvennogo-intellekta-v-pomoshch-stalevaru/](https://lipetsk.nlmk.com/ru/media-center/press-releases/nlmk-vnedril-sistemu-iskusstvennogo-intellekta-v-pomoshch-stalevaru/) от 19.02.2020

На Новолипецком металлургическом комбинате, флагманской площадке Группы НЛМК, разработали и внедрили цифровой сервис, который помогает максимально точно дозировать ферросплавы при выплавке стали.

Ферросплавы – металлические добавки в расплав при производстве жидкой стали для придания ей необходимых свойств. Например, хром делает сталь нержавеющей, ванадий повышает ее прочность, молибден дает ей стойкость к высоким температурам. На НЛМК могут выпускать несколько сотен марок стали для каждой из которой нужно добавить строгое количество ферросплавов определенного вида.

Ранее при расчете количества и состава ферросплавов сталевар использовал инструкции и личный опыт. Теперь ему помогает сервис, который на основе анализа всего массива данных о прошедших плавках с высокой точностью прогнозирует оптимальный расход ферросплавов для получения целевого химического состава стали. Чтобы сделать это максимально эффективно, сервис обращается к базе данных предприятия за информацией о количестве имеющихся на складе ферросплавов, их взаимозаменяемости и стоимости. Машинный анализ этой информации помогает рассчитать оптимальный состав и точный объем ферросплавов для получения стали требуемого химсостава. Результаты расчетов выводятся на экран монитора в качестве рекомендации, и сталевару остается принять окончательное решение.

Систему внедрили в первом конвертерном цехе на обычном марочном сортаменте. После расширения сервиса на весь сортамент двух цехов сталеплавильного производства ожидаемый экономический эффект составит около 100 млн рублей в год.

**Информация о Новолипецком металлургическом комбинате**

Новолипецкий металлургический комбинат – крупнейший в России производитель стали, основная производственная площадка Группы НЛМК. Комбинат является ядром единой международной производственной цепочки Группы НЛМК с активами в России, ЕС и США.

Объем производства стали на Липецкой площадке составляет около 18% всего производства стали в России и около 80% всей стальной продукции Группы НЛМК.

Высококачественная металлопродукция НЛМК применяется в различных стратегически

важных отраслях экономики, от строительства и машиностроения до производства энергетического оборудования и труб большого диаметра.

По итогам 2019 года объем производства стали на основной производственной площадке Группы НЛМК составил 12,01 млн тонн.

# Видео

Искусственный интеллект в помощь сталевару:



<https://youtu.be/h-wWJ2CvuMo>

# Описание кейса

Источник: Для эффективной цифровизации постройте «озеро». Опыт металлургов. РБК: Цифровая Россия.

[http://digital-russia.rbc.ru/articles/dlya-effektivnoy-tsifrovizatsii-postroyte-ozero-opyt-](http://digital-russia.rbc.ru/articles/dlya-effektivnoy-tsifrovizatsii-postroyte-ozero-opyt-metallurgov/) [metallurgov/](http://digital-russia.rbc.ru/articles/dlya-effektivnoy-tsifrovizatsii-postroyte-ozero-opyt-metallurgov/)

Мировая промышленность, в целом, и металлургия, в частности, переживают цифровую трансформацию. Развитие конкурентоспособного металлургического предприятия трудно представить без систем работы с данными и средств машинного обучения. И российский бизнес в этой сфере не остается в стороне.



Современная металлургия — это отрасль с высокой степенью автоматизации, где используются сложные системы. Одно из наиболее эффективных направлений в цифровизации на металлургическом производстве — комплексная работа с данными. В технологическом процессе производства их масса: это информация о химическом составе стали, температуре заготовки, давлении валков на полосу металла, режимах охлаждения и многом другом. В современных цехах используется множество систем и датчиков, непрерывно собирающих и передающих данные о производственных процессах. И эти данные можно использовать для оптимизации производства.

Работу с данными в широком смысле можно назвать термином big data. Так называют совокупность инструментов, подходов и методов обработки огромного объема данных для их последующего использования. Одним из важнейших подразделов big data являются средства машинного обучения (machine learning, ML) и искусственного интеллекта (ИИ). Машинное обучение — это методы, по которым в том числе ИИ способен самостоятельно учиться на примерах и ошибках, извлекая закономерности и делая выводы. ML-средства показывают лучшие результаты при работе с большими данными, помогают сократить издержки и оптимизировать расход времени и сил. В металлургии они могут, например, помочь оптимизировать расход сырья и добавок при производстве стали.

Значение машинного обучения в современной промышленности переоценить трудно. По данным консалтингового агентства Gartner, объем глобального рынка искусственного интеллекта к концу 2018 года составил $1,2 трлн, а к 2021-му достигнет отметки $2,9 трлн.

# Металлургический гигант устроил цифровую трансформацию

Несмотря на относительно низкий уровень проникновения big data, а также ИИ- и ML-технологий в нашей стране, российские компании не игнорируют эти направления.

Подобные решения активно внедряет группа НЛМК. Это настоящий гигант отрасли: НЛМК производит 17 млн т стали в год, а показатель EBITDA по итогам 2018-го составил $3,6 млрд. Его акции размещены на Московской и Лондонской фондовых биржах.



В условиях глобальной конкуренции на мировом рынке стали перед НЛМК стоят те же вызовы, что и перед другими компаниями отрасли. А именно: развитие на рынках в условиях роста протекционизма и избыточных мощностей, повышение эффективности производства, снижение воздействия на окружающую среду, повышение безопасности производства, развитие человеческого капитала.

Для повышения конкурентоспособности на мировом рынке компания реализует масштабную программу по улучшению операционной эффективности, это в том числе помогает снизить себестоимость продукции. В условиях сложных процессов на оборудовании инструменты машинного обучения способны дать существенную экономию. Возможности для этого есть на каждом этапе создания стоимости продукции. Помочь здесь способны комплексная работа с данными и применение ML-инициатив, для чего требуется удобный инструмент для хранения и обработки данных с производственных систем.

На НЛМК разработали стратегию работы с данными и построили кластер для централизованного накопления данных с датчиков и систем управления производством. В сентябре 2019-го в компании началось внедрение на производстве системы анализа данных и моделирования (САДиМ). Это платформа класса решений data lake, которая позволяет сохранять и предоставлять информацию о производственных и технологических процессах моделям машинного обучения. Также она помогает в разработке других цифровых сервисов и оптимизации процессов.

Техническая команда НЛМК совместно со специалистами «Инфосистем Джет» создали [«озеро данных» (data lake)](https://jet.su/datalake/?utm_source=rbc&utm_medium=article&utm_campaign=promo&utm_content=landing_page) — горизонтально масштабируемое хранилище больших объемов необработанной информации. Это один из ключевых элементов системы big data.

Евгений Колесников – руководитель Центра машинного обучения

«Инфосистем Джет»:

«На сегодняшний день главным потребителем решений, аналогичных САДиМ, является любое тяжелое производство, в частности черная, цветная металлургия. Почему эти решения интересны и

востребованны? Потому что металлургия сама по себе — это длинная производственная цепочка. На каждом участке производства создается добавленная стоимость. Оптимизируя даже по 3–5% на каждом этапе, за счет их большого количества на выходе заказчик может получить существенную выгоду. Именно это и является ключевым фактором для внедрения систем такого класса».

Важно понимать, чем data lake отличается от традиционных хранилищ. Данные в

«озере» сохраняются в оригинальном формате, они доступны всем доверенным пользователям в компании. Кроме того, data lake существенно дешевле классических хранилищ, так как для его эффективной работы требуется менее дорогостоящее оборудование и не требуется длительное приведение данных к единой модели. «Озеру» нужны лишь относительно дешевые сервера, а распределенная файловая система позволяет в любой момент подключать новые узлы без изменения структуры данных. Потому data lake может масштабироваться до экстремальных объемов при низких затратах. «Озера» и традиционные хранилища данных (data warehouse, DWH) не заменяют, а дополняют друг друга. Data lake дает компании новые возможности, позволяя хранить большие объемы данных и избежать затяжного процесса структурирования информации.

Евгений Колесников – Руководитель Центра машинного обучения «Инфосистем Джет» «У роста объема big data есть причины: предприятия начинают видеть цену этих данных. Их основным источником являются различные сенсоры, которые отдают данные очень часто, несколько раз в секунду. Затем эти данные аккумулируются и усредняются, например, записывается средний показатель раз в минуту. Чаще всего эти усреднения являются довольно грубыми — их недостаточно. И заказчики начинают записывать данные чаще: два раза в минуту, десять раз в минуту, каждую секунду — и в итоге доходят до

«сырых» данных. Для их хранения требуется очень большой объем, потому и используются

решения вроде САДиМ. Когда эти «сырые» данные забираются в неагрегированном виде прямо из источника, это дает прирост с точки зрения точности модели. ***А каждый процент в этой модели — это конкретные деньги».***

Специалисты настроили регулярную загрузку данных в «озеро» из 70 источников, а также загрузили исторические данные за последние несколько лет работы предприятия и разработали карты данных технологических и производственных процессов отдельных цехов. По просьбе НЛМК специалисты «Инфосистем Джет» разработали унифицированную витрину данных, реализовали контроль метаданных с помощью набора базовых сервисов управления Apache Atlas и настроили централизованную ролевую модель, интегрировав ее со службой Active Directory.



Дмитрий Кулагин – Руководитель направления data lake компании

«Инфосистемы Джет» «Сама идея платформы САДиМ принадлежит экспертам НЛМК, а наши специалисты проводили технологические и технические работы непосредственно по реализации данного решения на платформе Arenadata Hadoop. НЛМК выбрал нашу компанию, так как

мы предложили оптимальную архитектуру потоков загрузки и описания данных, и к тому же мы обладаем высоким уровнем экспертизы по Hadoop и интеграционным решениям, подтвержденным сертификатами наших архитекторов, разработчиков и системных администраторов. Кроме того, в нашей команде есть опытные аналитики, которые построили не одно хранилище данных, а также высококлассные тестировщики и специалисты по CI/CD».

Проект data lake в НЛМК был представлен на [RAIF 2019](https://raif.jet.su/) — бизнес-форуме по системам искусственного интеллекта, который прошел 22–23 октября в Сколково.

# Как выстраивается комплексная работа с данными. Этапы

Первые пилотные проекты. На этом этапе происходит построение пилотных моделей машинного обучения, которые могут быть как рекомендательными, так и оценочными. В ходе выполнения пилотов даются рекомендации по целесообразности применения моделей ML для данной задачи и возможные пути повышения качества предложенных моделей. Они оцениваются с точки зрения потенциального экономического эффекта и сложности реализации. Также определяется порядок приоритетности задач. При разработке пилотов уточняются требования к будущей экосистеме данных и моделей, осуществляется погружение специалистов по данным и ИТ в специфику производства, само производство знакомится с новыми инструментами.

Создание стратегии data science. Важно понимать, как именно управлять жизненным циклом данных, как поддерживать корпоративную модель данных. Надо понимать, какие данные есть в компании, чтобы разработать правильную стратегию работы с ними. Тогда в ней будет ценность для бизнеса. Параллельно с первыми пилотами вырабатывается и уточняется «дорожная карта» по созданию платформы для data science, включающая как развитие платформы хранения, так и подходы к работе с ML-моделями.

Внедрение data lake как один из первых шагов. Более дешевые и функциональные по сравнению с традиционными хранилищами «озера» позволяют быстро обрабатывать данные с помощью аналитики и машинного обучения. Их концепция позволяет начать накапливать данные еще до определения конкретных задач. В свою очередь, это дает возможность использовать исторические данные для ML-моделей. Развертывание слоя хранения данных и загрузка туда доступной истории приводит к повышению скорости апробации и внедрения новых моделей.

Накопление данных. Оно идет в любом случае, даже если данные не используются в существующих моделях. Важно организовывать пространство для хранения и заниматься минимальным структурированием, иначе «озеро» превратится в бесполезное «болото». Кроме того, необходимо связать data lake с аналитической экосистемой компании и обеспечить безопасность информации: она не должна «утекать» или создавать проблемы с регуляторами.

Создание слоя моделей и вывод их в продуктив. Для использования «озера» не нужно ждать долгого накопления и трудоемкой структуризации данных. Использование data lake и современных технологий виртуализации позволяет в кратчайшие сроки перейти к развертыванию слоя для моделей и их созданию в целевой архитектуре. Со временем

технологии и состав данных меняются, качество модели может падать, из-за чего нужно ее модифицировать либо создавать новую. Со временем может возникнуть несколько моделей, которые в разных ситуациях могут быть более или менее эффективны. Поэтому одновременно с разработкой моделей-сервисов создаются инструменты, которые позволяют управлять их жизненным циклом.

# Будущее за технологией data lake и ML-инициативами

Комплексная работа с данными, ML-инициативы и цифровизация процессов позволяют любой производственной компании стать эффективнее. Это дает возможность максимизировать прибыль за счет снижения себестоимости продукции, облегчить и ускорить работу специалистов, повысить безопасность производства, улучшить ситуацию с перерасходом сырья, процентом брака и обслуживанием оборудования. А в перспективе

* обеспечить переход к полностью автономному производству.

Технологии data lake и машинного обучения нужны не только в металлургии. Их применяют в нефтегазовой, финансовой, страховой, транспортной, маркетинговой и других отраслях, а также в ритейле и e-commerce.

Рост эффективности компании после их внедрения дает массу плюсов: от роста стоимости акций компании до повышения привлекательности в глазах потенциальных инвесторов. И сотрудничать в этой области стоит с надежным партнером. Ведь качественная работа с данными — залог успешного функционирования компании в XXI веке.

# Дополнительная информация:

* 1. НЛМК: искусственный интеллект на службе у сталеваров// Директор информационной службы. 2018. № 1. <https://www.osp.ru/cio/2018/01/13053828>
  2. Смирнов Н. Искусственный интеллект на НЛМК: конвейер оптимизации. Вестник цифровой трансформации CIO.RU 20.02.2019.

[https://www.computerworld.ru/cio/news/200219-Iskusstvennyy-intellekt-na-NLMK-](https://www.computerworld.ru/cio/news/200219-Iskusstvennyy-intellekt-na-NLMK-konveyer-optimizatsii) [konveyer-optimizatsii](https://www.computerworld.ru/cio/news/200219-Iskusstvennyy-intellekt-na-NLMK-konveyer-optimizatsii)