



**КЛАСТЕК**  
Аналитика Big Data

## **ИСПЦ ОРЭМ**

**Интеллектуальная система  
прогнозирования цен на ОРЭМ  
для генерирующих предприятий,  
энергосбытовых компаний и  
крупных энергопотребителей**

## Содержание

|  |   |
|--|---|
| <b>Назначение системы</b>                                    | 3 |
| <b>Задача прогнозирования энергопотребления</b>              | 4 |
| Недостатки классических методов прогнозирования              | 4 |
| <b>ИСПЦ – прогнозная модель на основе машинного обучения</b> | 5 |
| Преимущества машинного обучения                              | 5 |
| Функции системы  | 5 |
| <b>Технические характеристики системы</b>                    | 6 |
| Используемые алгоритмы                                       | 6 |
| Почему алгоритмы показывают хороший результат                | 6 |
| Проверка модели  | 7 |
| <b>Использование модели участниками ОРЭМ</b>                 | 8 |
| <b>Наши контакты</b>   | 9 |

## Назначение системы

Интеллектуальная система прогнозирования цен на электроэнергию (ИСПЦ) предназначена для использования на оптовом рынке электроэнергии и мощности (ОРЭМ) для рынка на сутки вперед (РСВ).

ИСПЦ позволяет формировать краткосрочные и среднесрочные прогнозы с учетом множества входных параметров и меняющихся внешних факторов. Основу системы составляет искусственный интеллект, построенный на алгоритмах машинного обучения и анализа больших данных.

### Функции системы

- формирование прогноза равновесной цены на РСВ;
- выявление факторов, оказывающих наибольшее влияние на результат прогноза.

Система предназначена как для генерирующих компаний, так и для энергосбытовых компаний и крупных энергопотребителей.

### Преимущества прогнозирования будущих ценовых показателей РСВ

#### Потребление

- ✓ Рациональная закупка необходимых объемов электроэнергии

#### Генерация

- ✓ Оптимизация состава и параметров генерирующего оборудования

#### Потребление и генерация

- ✓ Снижение убытков при выходе на балансирующий рынок

ИСПЦ может быть внедрена в ИТ-инфраструктуру заказчика и использоваться в качестве дополнительного инструмента по выбору оптимальной стратегии на РСВ.



Визуализация прогноза ИСПЦ

## Задача прогнозирования энергопотребления

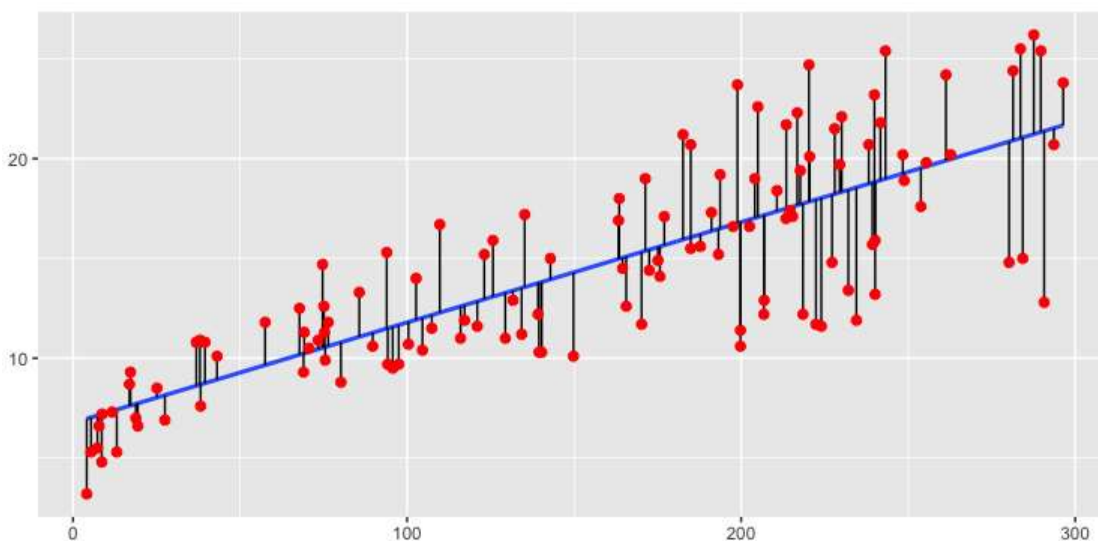
Для энергопотребителей - участников ОРЭМ наиболее актуальны кратко- и среднесрочные прогнозы равновесной цены на электроэнергию на РСВ. Точное прогнозирование будущих ценовых показателей РСВ позволяет рационально закупать необходимые объемы электроэнергии и снизить до минимума убытки при выходе на балансирующий рынок.

### Недостатки классических методов прогнозирования

Для решения задач прогнозирования равновесной цены РСВ обычно используются линейные регрессионные модели, основанные на анализе временных рядов. Основным недостатком линейных моделей является то, что они описывают именно линейную зависимость между влияющими факторами и результатом, в то время как прогнозирование рынка электроэнергии приводит к более сложным зависимостям.

Другим существенным недостатком линейных моделей является чувствительность к шумам в данных. Точность линейной модели серьезно зависит от обработки выбросов и пропущенных значений. Также успешность применения моделей линейной регрессии сильно зависит от выполнения предпосылок Гаусса-Маркова, которые на практике за редким исключением не выполняются.

Все эти недостатки снижают качество прогнозирования, а значит увеличивают расходы на покупку электроэнергии на ОРЭМ.



Классическая линейная регрессионная модель

## **ИСПЦ – прогнозная модель на основе машинного обучения**

Интеллектуальная система прогнозирования на ОРЭМ, разработанная компанией “Кластек”, позволяет формировать краткосрочные и среднесрочные прогнозы цены с учетом меняющихся внешних факторов и нелинейных закономерностей.

Основу системы составляет прогнозная модель, построенная на алгоритмах машинного обучения и анализа больших данных.

### **Преимущества машинного обучения**

Основная привлекательность применения алгоритмов машинного обучения для прогноза состоит в возможности использования большого количества входных параметров модели, которые повышают точность итогового прогноза.

Помимо стандартных параметров для энергетического рынка, таких как прогноз потребления, прогноз генерации, факт потребления, факт генерации, в модели могут учитываться следующие параметры:

- нагрузки за предшествующее время;
- погодные условия (температура, скорость ветра, облачность и т.д.);
- время суток;
- тип дня (рабочий, выходной, праздничный), день недели.

Методы машинного обучения позволяют обнаруживать и анализировать сложные нелинейные закономерности между использованными параметрами и выделять те, что оказывают наибольшее влияние на прогноз.

### **Функции системы**

- формирование прогноза на целевую переменную - равновесную цену на РСВ;
- анализ входного потока данных и выявление предикторов, оказывающих наибольшее влияние на формирование целевой переменной.

Точность модели повышается за счет следующих особенностей ее архитектуры:

- включение предыдущих наблюдений как признаков, а не как отдельных моделей;
- денормализация весов для более устаревших наблюдений;
- кросс-валидация для проверки точности прогноза.

После предъявления входных параметров, возможно вместе с желаемым выходом, модель самонастраивается (обучается) для обеспечения требуемой реакции. Для обучения модели используется алгоритм, цель которого состоит в уменьшении ошибки между выходом и требуемым значением, выбранным для обучения.

## Технические характеристики системы

### Используемые алгоритмы

В основе прогнозной модели лежат следующие алгоритмы машинного обучения:

- рекуррентные нейронные сети;
- градиентный бустинг над решающими деревьями.

Выбор алгоритмов обусловлен их лучшим результатом по сравнению с другими методами для точечного прогноза, как на вариационном, так и на тестовом наборе данных. При этом среднее значение прогноза на различных валидационных выборках показывает низкий разброс.

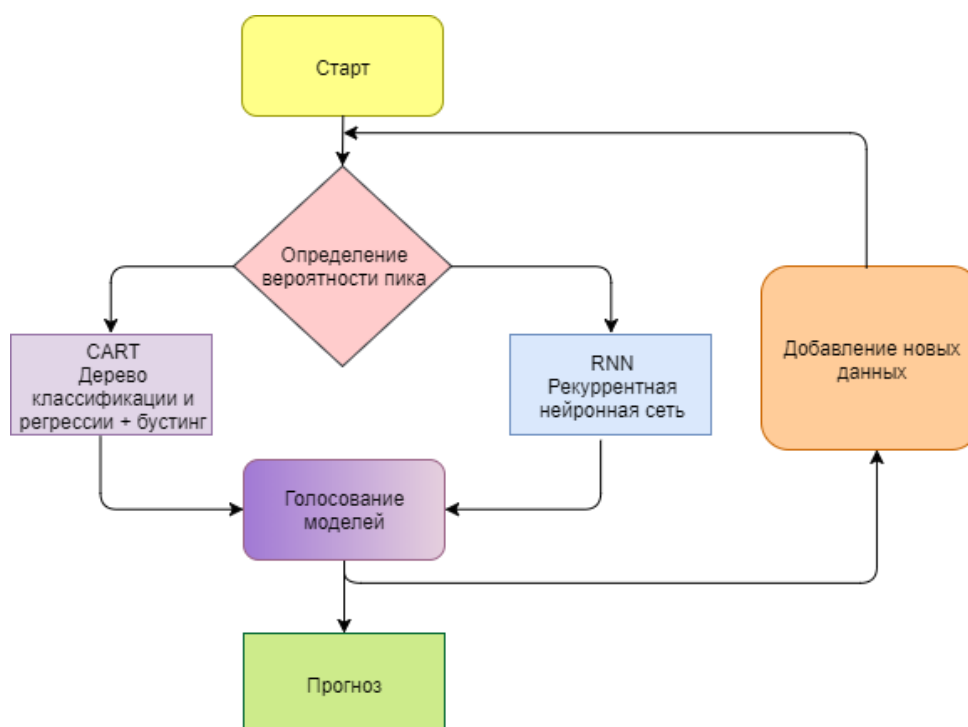
### Почему алгоритмы показывают хороший результат

**Рекуррентные нейронные сети** способны отражать сложные зависимости и улавливать повторяющиеся закономерности (паттерны) в данных благодаря своей архитектуре, которая представляет собой несколько слоев нейронов, причем нейроны каждого последующего слоя являются функцией от выходов предыдущего слоя, которые тоже в свою очередь являются функцией.

Кроме того, рекуррентные сети используют значения слоев, полученные на предыдущих наблюдениях. Такая методика хорошо работает на временных рядах, где последующие наблюдения сильно коррелированы с предыдущими.

**Градиентный бустинг над деревьями** работает эффективно вследствие того, что искомая зависимость между целевой переменной и объясняющими переменными строится в виде кусочно-постоянной функции. Это позволяет выявлять зависимости произвольной формы.

Кроме того, модели данного класса не чувствительны к выбросам в данных и имеют встроенные эвристики для обработки пропущенных значений.



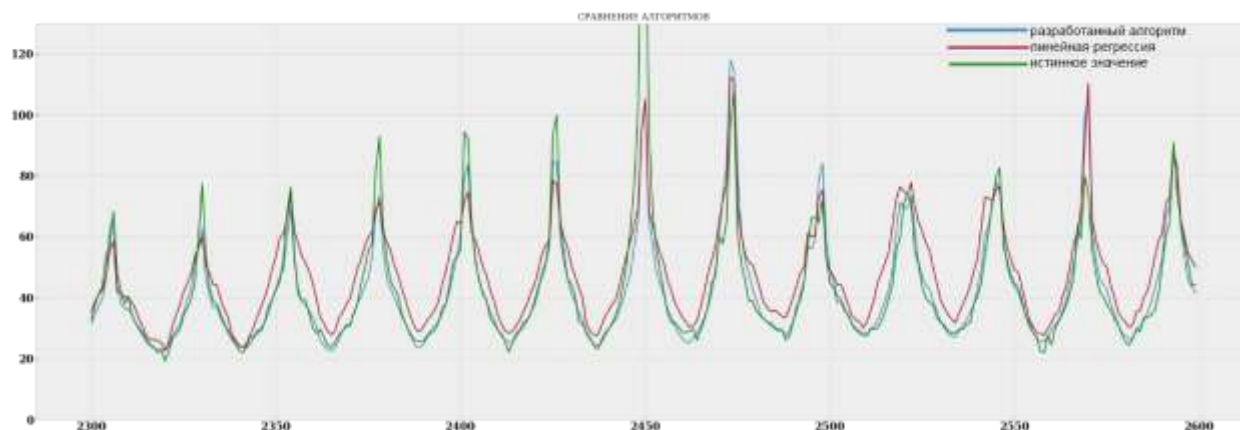
Архитектура прогнозной модели ИСПЦ

## Проверка модели

Разработанная модель была протестирована на данных по ценам рынка электроэнергии штата Техас, США (ERCOT) с учетом локальной специфики.

Был проведен анализ исторических данных рынка ERCOT в совокупности с данными метеостанций в конкретном регионе за последние 5 лет, который позволил построить оптимальную модель машинного обучения для прогнозирования цены на электричество как на рынке «на сутки вперед», так и на балансирующем рынке.

Разработанная модель показала лучшие результаты чем традиционный метод линейной регрессии. Результаты сравнения приведены на графике ниже.



На приведенном выше графике по оси Y расположены ценовые значения единицы энергии МВт/ч на Техасской бирже электроэнергии, торгуемые на рынке на сутки вперед, а по оси X – номера наблюдений. Зеленым цветом обозначено фактическое значение цены на рынке, красным и синим – модель линейной регрессии и разработанный алгоритм соответственно.

Из графика видно, что наш алгоритм позволяет аппроксимировать кривую фактических значений точнее, чем используемая в реальном производстве модель линейной регрессии.

Для достоверного сравнения точности аппроксимаций линейной регрессии и разработанной нами модели (в соответствии с графиком выше) взята распространенная метрика MAE (Mean Absolute Error – средняя абсолютная ошибка).

| Модель                      | MAE          | В процентах от среднего |
|-----------------------------|--------------|-------------------------|
| Линейная регрессия          | 7.165        | 17.8                    |
| <b>Разработанная модель</b> | <b>4.382</b> | <b>10.9</b>             |

Разработанная модель показывает значительно меньшую ошибку прогноза, чем линейная регрессия. При усечении верхней границы целевой переменной, в данном случае – цены МВт/ч, по величине, равной среднее + 3 стандартных отклонения, мы можем получить еще более точные результаты.

| Модель (при усечении целевого значения) | MAE          | В процентах от среднего |
|---|--------------|-------------------------|
| Линейная регрессия                      | 6.146        | 15.7                    |
| <b>Разработанная модель</b>             | <b>3.488</b> | <b>8.9</b>              |

## Использование модели участниками ОРЭМ

Схема американского и российского спотового рынка электроэнергии основана на идентичной узловой модели рынка, но рынок электроэнергии США является более волатильным, а потому более сложным для прогнозирования, чем рынок электроэнергии РФ.

Учитывая, что для рынка США модель ИСПЦ обеспечивает необходимое качество прогнозирования, наша модель может быть с успехом адаптирована к российскому рынку ОРЭМ при условии получения релевантных данных от заказчика.

Для разработки задачи для конкретного узла от заказчика необходимо получить архив наблюдений, включающий следующие параметры:

- Дата
- Интервал, ч.
- Равновесная цена на РСВ, руб./МВт.ч.
- Объем покупки/продажи по регулируемым договорам, МВт.ч.
- Объем покупки/продажи на РСВ, МВт.ч.
- Объем покупки/продажи в обеспечение РД, МВт.ч.
- Минимальная/максимальная заявленная цена покупки/продажи, руб./МВт.ч.
- Равновесная цена на БР, руб./МВт.ч.
- Объем планируемой генерации/потребления, МВт.ч.
- Объем фактической генерации/потребления, МВт.ч.
- Требования заказчика к интеграции с существующей ИТ-инфраструктурой

В зависимости от объема входных параметров и степени неопределенности тех или иных факторов, влияющих на нагрузку, для обучения модели могут потребоваться ретроспективные данные за периоды продолжительностью до нескольких лет.

Чем больше релевантных данных используется для обучения модели, тем более точным будет прогноз.

Большой объем выборки также позволяет сформировать точную градацию параметров, оказывающих наибольшее влияние на формирование прогноза.

ИСПЦ может быть внедрена в ИТ-инфраструктуру заказчика и использоваться в качестве дополнительного инструмента по выбору оптимальной стратегии формирования ценовых заявок на РСВ.



## Наши контакты

ООО "Кластек"

ОГРН 1057748269435, ИНН 7735514090

124498, г. Москва, г. Зеленоград, Георгиевский проспект, дом 5,  
стр.1

+7 (499) 460-68-30

[info@clustech.ru](mailto:info@clustech.ru)

