

# Анализ данных работы ГТУ в системе «Smart Diagnostics»

*Автор: Дарья Николаевна Немирович-Скратун, Ведущий инженер-математик Ctrl2GO Solutions*

Анализ данных работы ГТУ является неотъемлемой частью диагностики. Он позволяет контролировать текущее состояние, расследовать уже случившиеся инциденты, а также делать прогнозы. В зависимости от поставленной задачи способ анализа данных отличается. Данная статья посвящена обзору видов анализа данных работы ГТУ, а также систем, в которых он осуществляется. В статье подробно рассмотрим наиболее популярные способы анализа данных, их преимущества и недостатки.

Анализ данных — это процесс изучения, очистки, преобразования и моделирования данных с целью обнаружения полезной информации, подтверждения выводов и поддержки принятия решений. Анализ данных работы турбин позволяет обеспечить их эффективную и безопасную эксплуатацию. К основным методам анализа, которые применяются при анализе данных работы ГТУ, относятся: мониторинг технического состояния, регрессионный анализ, акустический анализ, машинное обучение, анализ

временных рядов.

Каждый из этих методов имеет свои преимущества и недостатки и применим для различных задач. Например, для прогнозирования обычно применяют регрессионный анализ [1]. А для выявления отклонения используют мониторинг, статистический анализ и также методы машинного обучения. Каждый из способов анализа имеет свою трудоемкость и требования к вычислительным мощностям.

Далее рассмотрим наиболее распространённые методы анализа данных работы ГТУ на примере системы «SmartDiagnostics» Ctrl2GO Solutions.

## **Мониторинг технического состояния**

Мониторинг технического состояния является наиболее доступным и понятным способом анализа данных работы ГТУ. Он может осуществляться на уровне станции или на уровне центра мониторинга.

В зависимости от уровня используются различные системы. При мониторинге на нижнем уровне используются SCADA (Supervisory Control And Data Acquisition — диспетчерское управление и сбор данных) системы. Они собирают, хранят данные и позволяют пользователю (оператору/эксперту) эти данные визуализировать. Такой тип мониторинга позволяет оперативно просматривать значения технологических параметров с точностью до миллисекунд. Однако, пользователь в таких системах анализирует данные в ручном режиме, что трудоемко. Также растет вероятность пропуска внештатной ситуации, поскольку оператор увидит информацию только

после срабатывания предупредительной или аварийной сигнализации при превышении технологических уставок.

В центры мониторинга данные поступают с разных станций по защищенному каналу связи. При этом передаются данные не со всех датчиков, а только с тех, которые могут идентифицировать дефект для превентивных мер. Центры мониторинга используют для анализа данных специальные системы – системы мониторинга и диагностики/прогностики. В дополнение к возможностям SCADA систем они позволяют строить различные типы моделей, настраивать логику для выявления дефектов, строить прогнозы.

Системы мониторинга обладают универсальной структурой: они собирают, обрабатывают и анализируют данные, сравнивая их с различными моделями — математическими, физическими или гибридными [2]. На основе этого сравнения система или специалист делает выводы о наличии аномалий в работе оборудования и формирует прогнозы. Качество диагностики в системах мониторинга зависит от качества сбора и обработки данных и заложенного в систему математического метода. В зависимости от метода модель будет иметь разную точность, которая влияет на качество выявления аномалий и точность прогноза. Точность или закладывается при расчете или же наоборот рассчитывается на основании обучающей выборки и последующей валидации модели.

Основными математическими методами, используемых в системах мониторинга, являются градиентный бустинг деревьев решений, нечеткие логики, моделирование на основе подобию, нейронные сети, метод главных компонент.

Зарубежные системы, такие как SiePA Siemens, SmartSignal GE и AVEVA

(Avantis PRiSM) Schneider Electric, традиционно имеют более развитую математическую базу, прописанные бизнес-процессы и широкий функционал. Однако в настоящее время на российском рынке они не представлены, что заставляет предприятия искать альтернативные отечественные решения или разрабатывать собственные системы мониторинга.

Отечественные системы, такие как «ПРАНА» РДС, CyberPhysics, SmartDiagnostics Ctrl2GO Solutions и «Anomaly monitor» Силовые машины, разрабатываются с учетом специфики и требований российского рынка. Они часто ориентируются на лучшие практики и решения, разработанные зарубежными производителями, стремясь соответствовать мировым стандартам качества и функциональности.

Рассмотрим мониторинг технического состояния на примере системы «SmartDiagnostics». Основным окном для анализа является окно «общая информация». Вид его представлен на рисунке 1. В левой части рабочего окна выводится информация по работе статистических моделей [3]. Здесь пользователь видит количество отклоняющихся параметров от их общего числа. Также видно разделение на узлы. Правее пользователь видит информацию по работе экспертных моделей. В правой части пользователю доступен журнал событий. В него записываются и выводятся сообщения по срабатыванию различных правил (экспертных и диагностических), срабатывания уставок, отклонений модели. Также в системе есть окно общего мониторинга, в котором собрана информация по работе всех агрегатов, подключенных к системе.

В системе «SmartDiagnostics» также доступен обычный анализ данных посредством просмотра графиков. При этом на него можно выводить данные

с разных агрегатов, что удобно при анализе работы энергоблока или сравнительном анализе работы однотипных турбин.

Важным инструментом в системе помимо статистического моделирования является система правил: диагностических и экспертных. Система диагностических правил позволяет на основании данных работы модели определять наличие дефекта или аномалии. При этом выдается вероятностная оценка наличия того или иного дефекта, что существенно помогает при принятии решений. Экспертные правила выдают диагноз это на основании данных телеметрии.

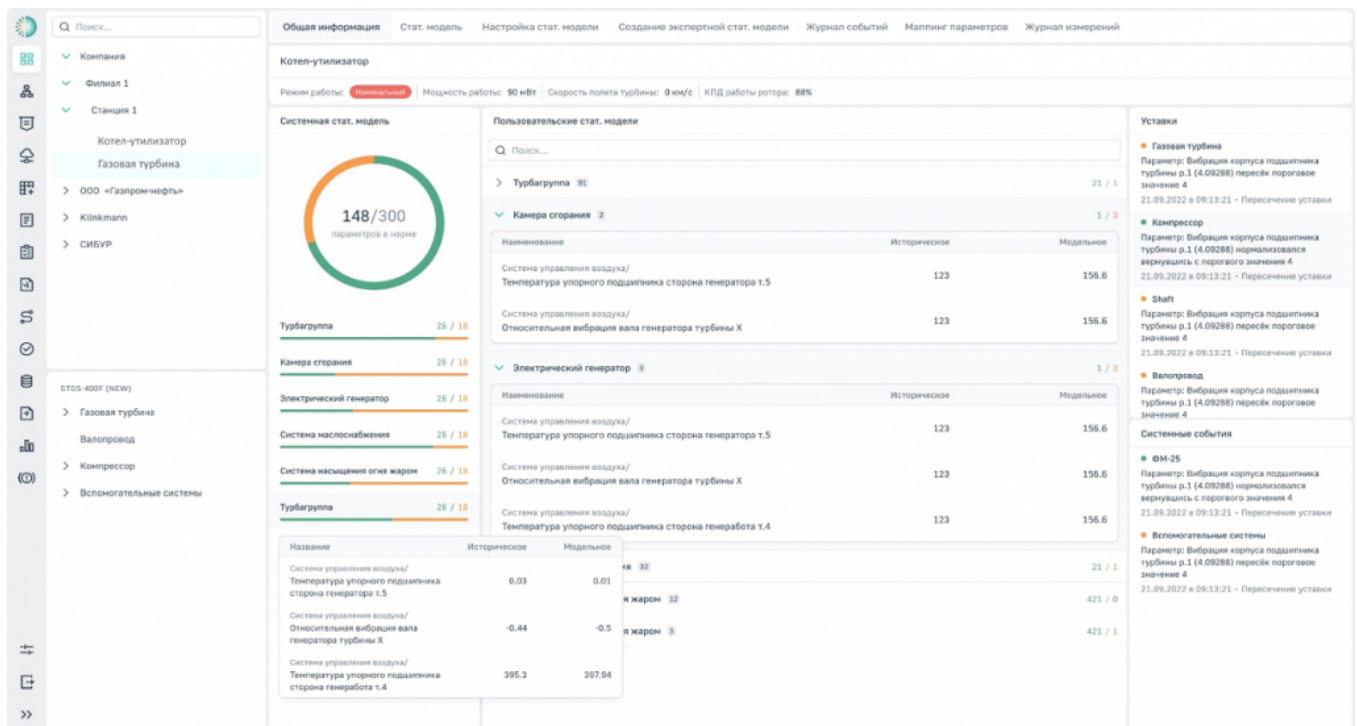


Рис.1 - мнемосхема мониторинга технического состояния

## **Машинное обучение**

Машинное обучение — это подраздел искусственного интеллекта, который фокусируется на разработке алгоритмов, способных учиться из данных и делать прогнозы или принимать решения на основе этих данных. В основе машинного обучения лежит идея, что системы могут учиться из опыта, адаптироваться к новым данным и выполнять задачи без явного программирования для каждого случая [4].

Машинное обучение делится на обучение с учителем, без учителя, с частичным привлечением. Также существует обучение с подкреплением. В случае обучения с учителем данные для обучения размечаются специалистом. При этом участие человека может быть частичным. В случае обучения без учителя данные не размечены, а алгоритм пытается найти в них закономерности. При обучении с подкреплением модель учится принимать решения, получая награды или наказания за свои действия в динамической среде.

Машинное обучение использует различные алгоритмы и математические методы, такие как линейная и логистическая регрессия, деревья решений, случайные леса, нейронные сети, и многие другие.

Наиболее популярным методов в машинном обучении являются нейронные сети. Они могут иметь различную структуру, каждая из которых применима к решению конкретных задач. Например, для поиска аномалий в данных о работе турбинного оборудования хорошо себя показывают рекуррентные сети. Их структура представлена на рисунке 2.

Для получения точной модели особенно важное значение имеет качество

подготовки данных, на основании которых будет строиться модель. Если в модель заложить заведомо некачественные данные, то модель не сможет выявлять отклонения или делать точные прогнозы.

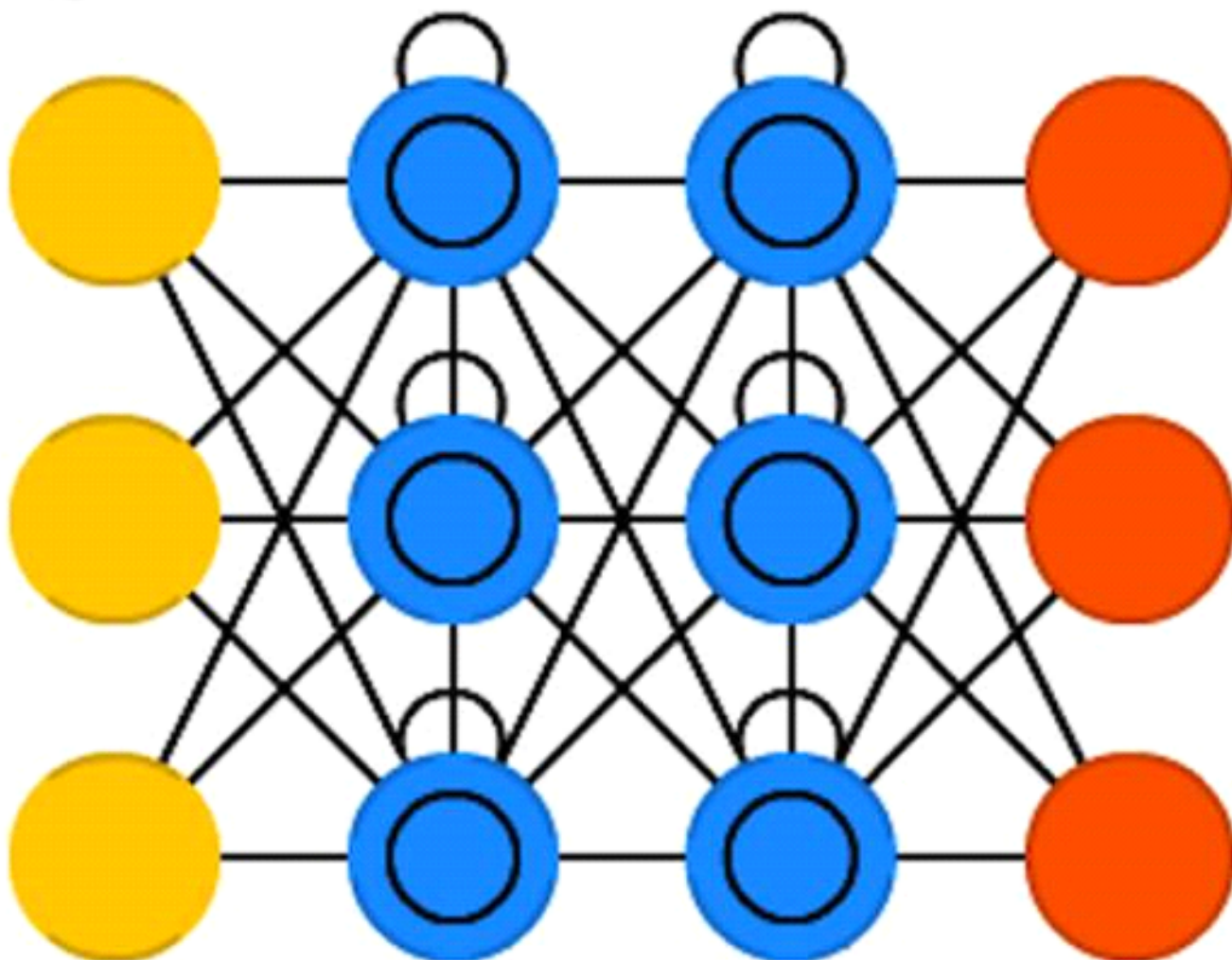


Рис.2 - схема рекуррентной нейронной сети

На данный момент некоторые системы мониторинга уже используют методы

машинного обучения. Например, решение CyberPhysics использует нейронные сети для выявления аномалий в работе оборудования. «SmartDiagnostics» использует регрессионные модели для прогнозирования и нейронные сети в качестве дополнительного метода анализа данных.

## **Регрессионный анализ**

Регрессионный анализ условно можно разделить на две части с точки зрения задач: выявление зависимостей между параметрами и прогноз технического состояния. В первом случае построение линейной регрессии позволяет построить корреляционную матрицу и провести анализ зависимостей между параметрами. Во втором случае регрессия позволяет посчитать прогноз отдельно взятого параметра. Это может быть любой технологический параметр или индекс технического состояния.

Чаще всего регрессионный анализ используют для прогнозирования технического состояния. Наиболее простой, но наименее точной является линейная регрессия [5]. Она встречается во всех современных системах мониторинга.

В основе регрессионной модели лежит процесс, который можно описать следующим уравнением:



$$Y(t) = \alpha_0 + \alpha_1 \times X(t) + \xi_t$$

$\alpha_0$  ,  $\alpha_1$  - коэффициенты модели;

$\xi_t$  - случайная ошибка.

Регрессионная модель может быть не только линейной, но и любой другой-полиномиальной, логарифмической.

Наиболее точной будет множественная регрессионная модель, которая оценивает связи между параметрами. При этом модель, как правило, линейна, поскольку она состоит из аддитивных членов, каждый из которых представляет собой предиктор, умноженный на оценку коэффициента. Как правило, в модель добавляется также константа (свободный член).

К регрессионным моделям относится также Метод группового учета аргументов (МГУА). Суть метода заключается в выборе регрессионных моделей оптимальной сложности. Чем сложнее модель, тем точнее прогноз. МГУА чаще всего применяется при наличии в выборке множества аргументов. В этом случае модель описывается уравнением:

$$Y(t) = w_0 = \sum_{i=1}^m w_i X_i(t) + \sum_{i=1}^m \sum_{j=1}^m w_{ij} X_i(t) X_j(t) + \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^m w_{ijk} X_i(t) X_j(t) X_k(t) + \dots \quad (2)$$

Где  $x = \{x_i = 1, \dots, m\}$  - множество свободных переменных;

$w_{ij}$  - вектор параметров весовых коэффициентов.

Уравнение (2) называется опорной функцией. Для всех или выборочных аргументов строятся модели. Каждая модель обладает весовым коэффициентом, из которых, на следующем шаге выбираются лучшие. Если качество найденных моделей и полученных из нее прогнозных значений удовлетворяет, то процесс прекращается. Иначе, на предыдущем шаге выбранные модели используются в качестве аргументов  $X_1(t) \dots X_n(t)$ .



Рис.3 – пример прогноза методом МГУА

Преимуществом регрессионных моделей являются их простота, быстрота получения результата, доступность промежуточных вычислений и разнородность решаемых задач, что делает эти модели применимыми во всех отраслях промышленности. При этом они имеют следующие недостатки:

- невысокая точность прогноза для интерполяционных данных;
- сложность определения вида функциональной зависимости;
- трудоемкость определения параметров;
- субъективный выбор модели (зависимости). Формально- подгонка под эмпирический материал.

## **Заключение**

Мониторинг технического состояния является наиболее простым и понятным способом анализа данных о работе ГТУ. При этом системы мониторинга помимо привычных инструментов (просмотр графиков) обладают более широкими возможностями. Они позволяют пользователю строить различные модели работы, делать прогнозы. Одной из таких систем является «SmartDiagnostics», которая представляет собой комплексное решение за

счет комбинации используемых в ней математических методов.

Для построения моделей в системе используют методы моделирования на основе подоби́я и нейронные сети. Они позволяют находить отклонения в работе турбомашин, детерминировать их и далее прогнозными методами планировать последующие действия. Для прогнозирования используется регрессионный анализ, а также регрессионные модели различной сложности.

Прогноз технического состояния турбомашин позволяет существенно снизить количество аварийных остановов, спланировать ремонтные работы и т.д. по этой причине наличие прогнозного блока в системах мониторинга является неотъемлемой составляющей. И именно этот блок представляет наибольший интерес с точки зрения пользователя.

## **Список литературы**

1. Свидетельство о государственной регистрации программы для ЭВМ № 2017619335 Российская Федерация. "Программный модуль "ПРАНА": автоматическая регрессия" : № 2017616260 : заявл. 28.06.2017 :опубл. 22.08.2017 / Д. Н. Скрабатун, Ю. В. Гиринский; заявитель Акционерное общество «РОТЕК».
2. Немирович-Скрабатун, Д. Н. Диагностика неисправностей работы энергетического оборудования с помощью системы SmartDiagnostics / Д. Н. Немирович-Скрабатун, А. К. Рыбаков // Автоматизация в промышленности. – 2022. – № 11. – С. 9-13.

3. Немирович-Скрабатун, Д. Н. Разработка алгоритма автоматического построения моделей работы энергетического оборудования в системе SmartMaintenance / Д. Н. Немирович-Скрабатун // Газотурбинные технологии. – 2022. – № 4(187). – С. 38-43.

4. Астапов, Р. Л. Автоматизация подбора параметров машинного обучения и обучение модели машинного обучения / Р. Л. Астапов, Р. М. Мухамадеева // Актуальные научные исследования в современном мире. – 2021. – № 5-2(73). – С. 34-37.

Айвазян, С. А. Прикладная статистика и основы эконометрики/ С. А. Айвазян, В. С. Мхитарян. – Москва: Юнити, 1998.