

УДК 621.9-1-9:004.896

**ОПИСАНИЕ ПРИМЕНЕНИЯ ПРОГРАММНОГО КОМПЛЕКСА
STATISTICA В ОБЛАСТИ МАШИННОГО ОБУЧЕНИЯ
ИСКУСТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ ПРИ ОЦЕНКЕ
ОСТАТОЧНОГО РЕСУРСА МАШИН И МЕХАНИЗМОВ**

Марков. И.В. , аспирант группы МТа-211, II курс
Научный руководитель – И. В. Чичерин, к.т.н., доцент
Кузбасский государственный технический университет имени Т.Ф.
Горбачева
Россия, г. Кемерово

Аннотация. Статья описывает применение программного комплекса STATISTICA для прогнозирования остаточного ресурса оборудования. Описаны основные возможности STATISTICA, а именно работа с нейронными сетями, анализ временных рядов и регрессионный анализ. Так же описаны основные этапы создания и обучения нейронных сетей в программе STATISTICA, а также способы их применения для повышения точности предсказаний отказов промышленного оборудования. В статье рассказывается о важности использования искусственных нейронных сетей для оптимизации процессов диагностики и обслуживания машин и механизмов, что помогает снизить затраты и повысить надежность.

Ключевые слова: STATISTICA, машинное обучение, искусственные нейронные сети (ИНС), прогнозирование остаточного ресурса, интеллектуальный анализ данных, нейронные сети, регрессионный анализ, анализ временных рядов, диагностика оборудования, предсказательная аналитика, оптимизация обслуживания, промышленные технологии.

1. Введение

В условиях современного производства, когда технологии быстро шагают вперед, одной из важнейших задач для большинства отраслей производства является предсказание остаточного ресурса оборудования. Непрогнозируемые поломки машин и механизмов влекут за собой большие финансовые потери и повышение рисков, которые являются нежелательными для любого предприятия. Традиционные методы диагностики и прогнозирования, зачастую не дают достаточной точности при планировании обслуживания, но тем не менее эти методы могут стать базой для разработки передовых методов оценки остаточного ресурса. При помощи новых методов, в том числе искусственных нейронных сетей (ИНС), у нас появляется возможность анализировать сложные взаимосвязи данных и точнее прогнозировать отказы и поломки в оборудовании.

Одним из мощных инструментов является программный комплекс STATISTICA. Данная программа предоставляет большой инструментарий для

обработки и анализа данных и временных рядов, а также дает возможность создавать и обучать нейронные сети. В статье рассмотрено применение программного комплекса STATISTICA для прогнозирования остаточного ресурса, что в свою очередь помогает повысить точность и надежность диагностики и прогнозирования.

2. Обзор возможностей STATISTICA для анализа данных

STATISTICA – универсальный инструмент для анализа данных, который включает в себя разные модули для решения задач из разных областей.

Основные модули программного комплекса STATISTICA:

- *Statistics или статистика*

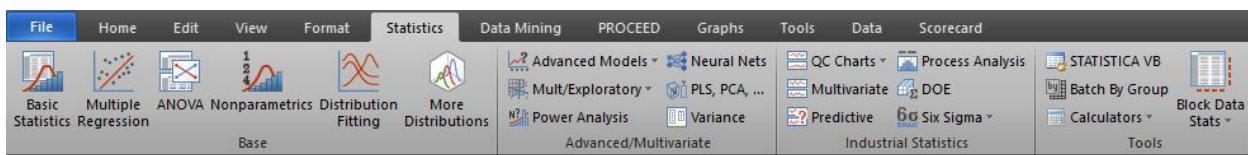


Рисунок 1. Панель "Statistics" в программе STATISTICA

Далее будет приведено краткое описание инструментов модуля Statistics:

1. *Basic statistics* – чаще всего называется «базовой статистикой» и представляет из себя группу инструментов, которые используются на начальном (исследовательском) этапе анализа данных. Включает такие инструменты как: описательная статистика, корреляция, временной анализ для независимых выборок, таблицы частот, перекрестные таблицы, дихотомия, тесты на разницу, калькулятор вероятности и др.

2. *Multiple Regression* – множественная регрессия, применяется для того, чтобы узнать больше о взаимосвязи между несколькими независимыми переменными и переменными-предикторами.

3. *ANOVA* – дисперсионный анализ, целью этого инструмента является проверка значимости различия между средними значениями.

4. *Nonparametrics* – непараметрическая статистика, этот инструмент применяется когда данные имеют ранжирование, но не имеют ясной численной интерпретации.

5. *Distribution Fitting* – подгонка распределений, используется когда есть сформулированная гипотеза о конкретном распределении интересующей нас переменной. К примеру, переменные, значения которых определяются бесчисленным числом независимых случайных событий требуют применения данной функции.

6. Так же присутствуют множество инструментов для продвинутого анализа статистических данных.

- *Data mining или интеллектуальный анализ данных*



Рисунок 2. Панель "Data mining" в программе STATISTICA

Краткое описание инструментов модуля Data mining:

1. *Data miner recipes* - Данный модуль позволяет, при помощи интуитивно понятного интерфейса, людям с малым опытом работы с данными проводить их анализ по принципу «Рецепта» который содержится внутри программы.

2. *Trees/partitioning* – набор инструментов, который дает возможность пристраивать «деревья» классификаций и регрессии для прогнозирования непрерывных зависимых переменных и категориальных переменных-предикторов. Такие методы «древовидной» классификации дают точные прогнозы или предсказания основанные на логических условиях «если – то».

3. *Learning* – инструменты для работы с нейронными сетями. В этом разделе присутствуют механизмы для создания и обучения нейронных сетей на базе исторических и статистических данных.

4. *Clustering/grouping* – набор инструментов для разделения и группировки данных с помощью анализа независимых компонентов. Эти методы используются для выделения интересующих нас данных из всей базы данных.

5. *Text mining* – модуль предоставляющий мощные инструменты для обработки неструктурированной информации, извлечения из текста нужных числовых показателей, это нужно для того, что бы сделать информацию находящуюся в тексте доступной для различных алгоритмов анализа данных.

6. *Deployment* - инструменты для генерации прогнозов на основе одной или нескольких обученных моделей.

3. Роль нейронных сетей в STATISTICA. Создание нейронной сети на базе программы STATISTICA

Одной из основных функций программного комплекса STATISTICA является создание и обучение нейронных сетей. Искусственные нейронные сети являются незаменимым инструментом, когда речь идет о работе с огромным количеством данных. Так же ИНС помогают приставать сложные взаимосвязи между данными, которые в свою очередь влияют на остаточный ресурс оборудования.

В программе STATISTICA за создание и обучение нейронных сетей отвечает модуль под названием Neural Networks. Данный модуль предлагает разные архитектуры нейронных сетей, которые помогают в решении самых различных задач – от классификации до регрессии и кластеризации.

Порядок создания нейронной сети в STATISTICA:

1. Первым шагом для создания и обучения нейронной сети ввод данных в таблицу программы STATISTICA.

2. После того как данные введены нужно вызвать инструмент «Neural Networks».

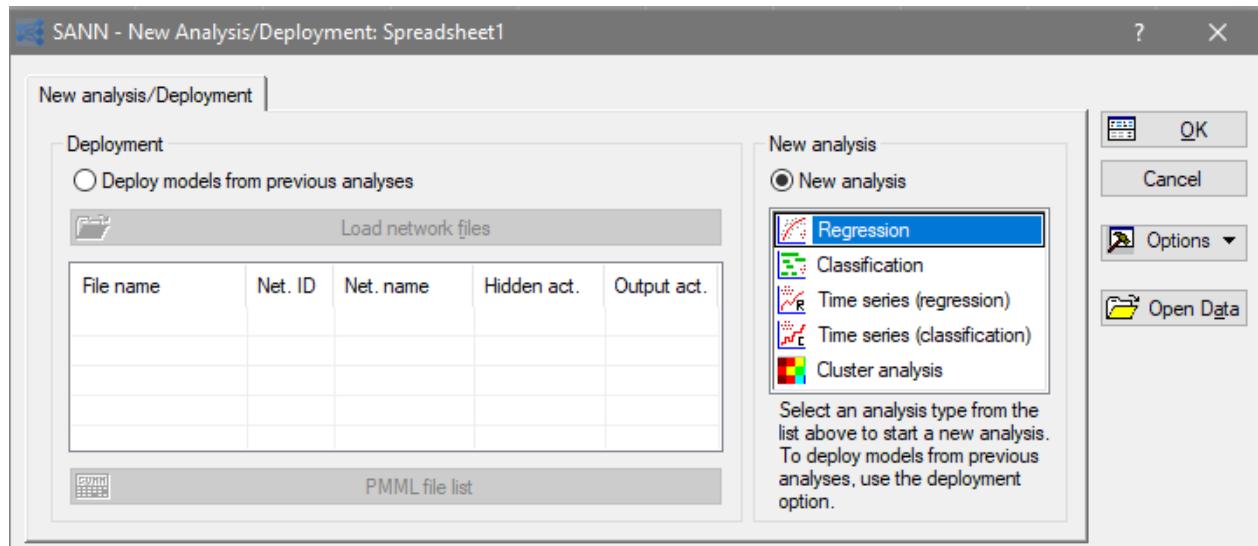


Рисунок 3. Окно "Neural Networks"

3. В правой части окна мы видим, что можно выбрать тип анализа для нейронной сети. После нажатия кнопки «OK» мы переходим к следующему окну. (рис.4)

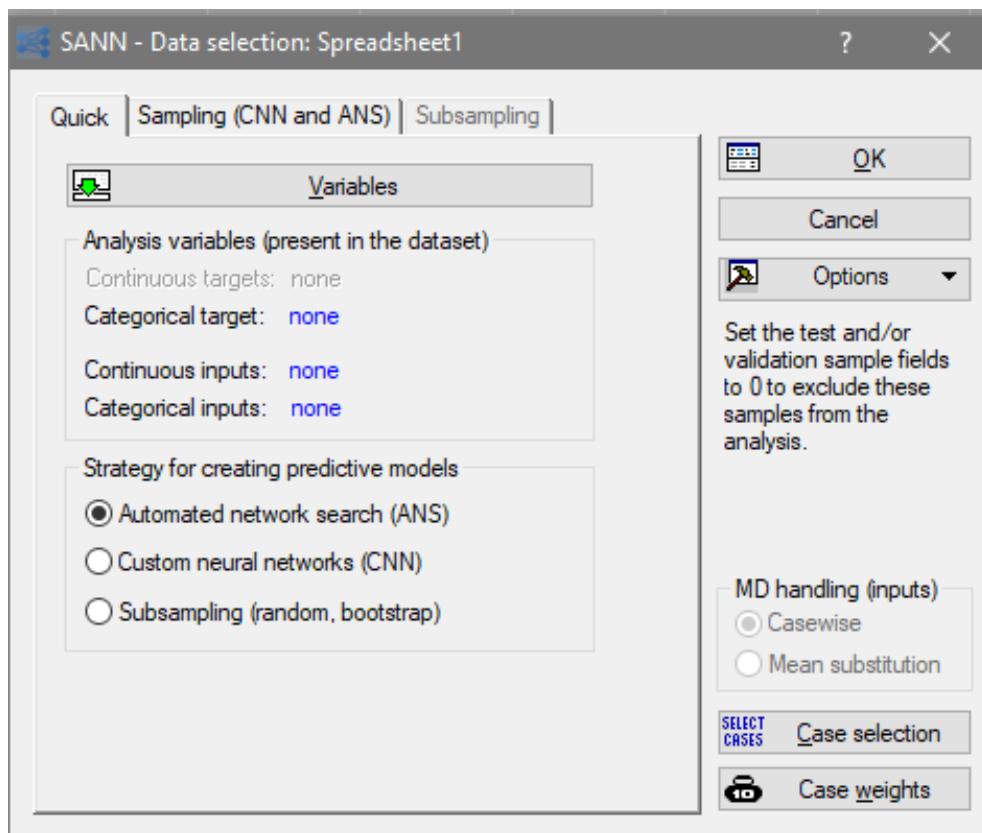


Рисунок 4. Окно "Data selection"

4. В выпавшем окне (рис. 4) мы можем выбирать стратегию для создания модели, в при нажатии кнопки “Variables” мы можем выбирать переменные для анализа.

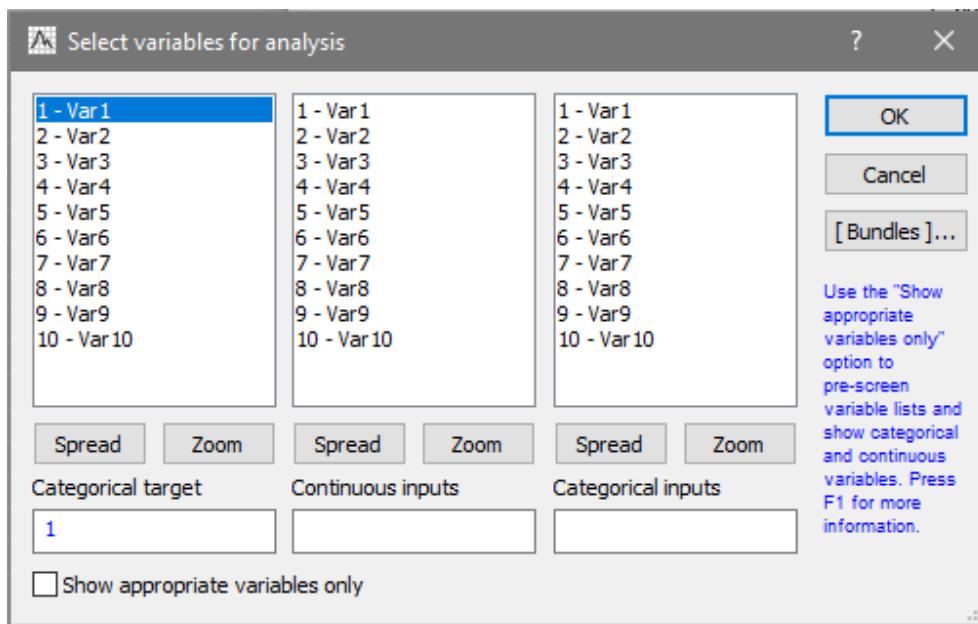


Рисунок 5.Окно выбора переменных для анализа.

5. После выбора переменных мы попадаем в окно выбора параметров обучения нейронных сетей.

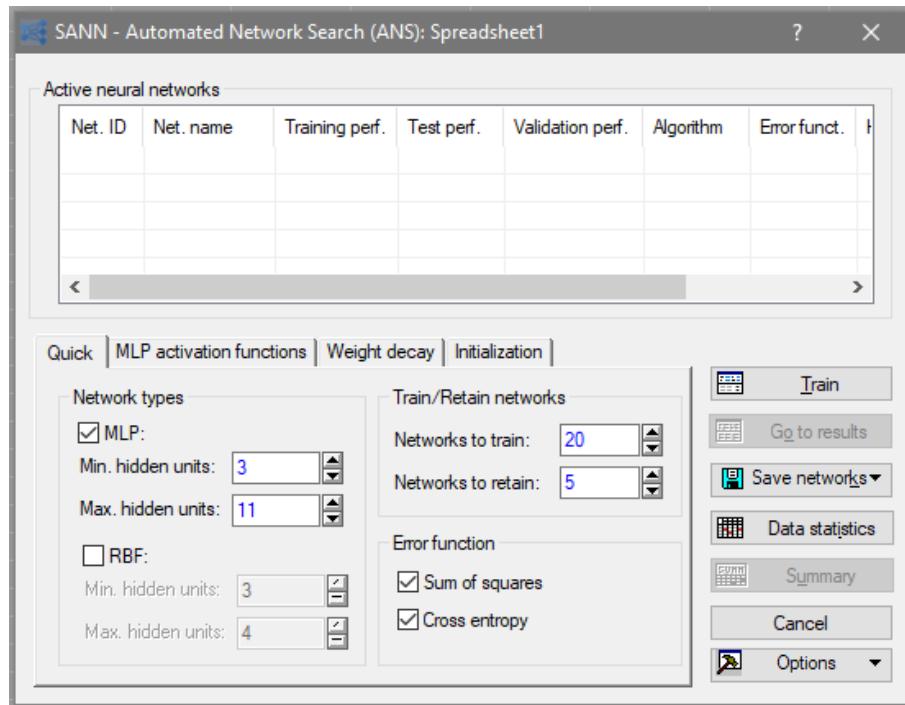


Рисунок 6. Окно выбора рабочих параметров обучения нейронной сети

6. После нажатия кнопки «Train» произойдёт обучение и построение нейронной сети с выбранными параметрами. После откроется окно для анализа построенной нейронной сети.

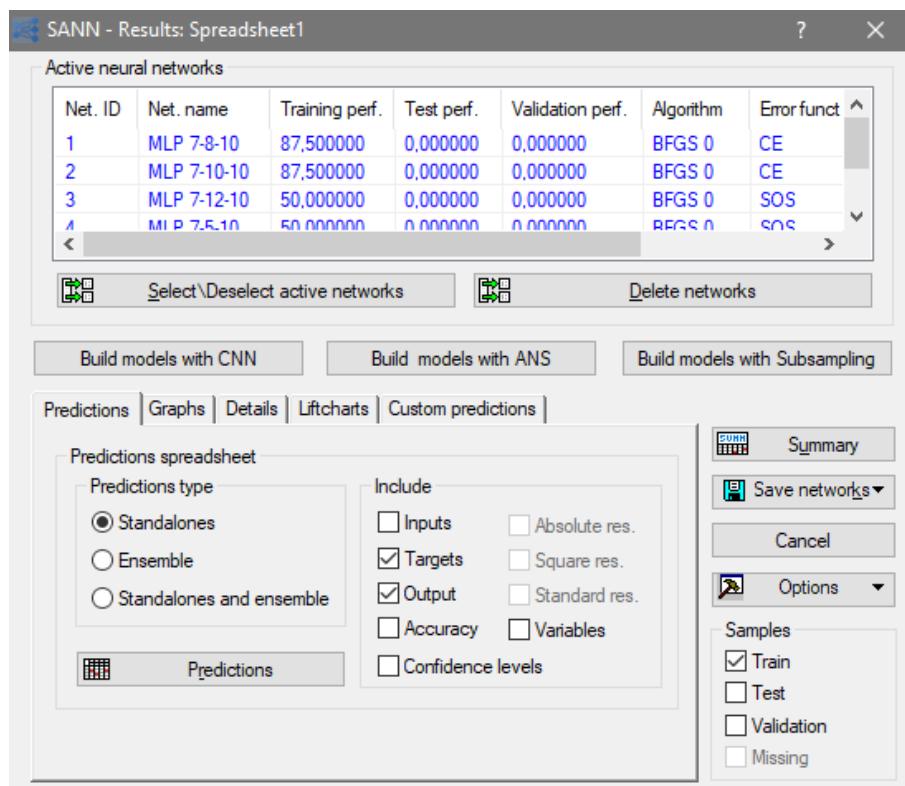


Рисунок 7. Окно результатов

7. Данные анализа автоматически выводятся в таблицу. Таким образом мы сможем понять какие параметры нашей нейронной сети наиболее значимы, то есть к каким параметрам наиболее чувствительна эта нейронная сеть.

8. После нужно проводить проверку адекватности построенной модели на реальном оборудовании.

В итоге рассмотрен общий алгоритм создания и обучения нейронной сети. Данные взяты условно для того, чтобы показать сам алгоритм.

ЗАКЛЮЧЕНИЕ

STATISTICA – мощный и удобный инструмент, который значительно облегчает процесс анализа данных, а так же дает возможность создавать и обучать нейронные сети. Программа может применяться в различных отраслях, а для её освоения вовсе не обязательно иметь глубокие познания в статистическом анализе.

Использование программы способно существенно повысить точность предсказания отказов машин и механизмов, что в свою очередь ведет к росту надежности оборудования и снижению затрат на его эксплуатацию.

Таким образом STATISTICA является актуальным и надежным инструментом для прогнозирования остаточного ресурса оборудования.

Список литературы:

1. Боровиков, В.П. Нейронные сети. STATISTICA Neural Networks: Методология и технологии современного анализа данных. – М.: Горячая линия. – Телеком, 2008. – 392с
2. Боровиков, В.П. Популярное введение в современный анализ данных в системе STATISTICA. Учебное пособие для вузов. – М.: Горячая линия – Телеком, 2013. – 288 с.
3. Воронцов В.А., Федоров Е.А. Разработка прототипа интеллектуальной системы оперативного мониторинга и технического состояния основных бортовых систем космического аппарата // ТРУДЫ МАИ. Учредители: Московский авиационный институт, 2015. - №82. -35с
4. ГОСТ 27.302-86. НАДЕЖНОСТЬ В ТЕХНИКЕ. МЕТОДЫ ОПРЕДЕЛЕНИЯ ДОПУСКАЕМОГО ОТКЛОНЕНИЯ ПАРАМЕТРА ТЕХНИЧЕСКОГО СОСТОЯНИЯ И ПРОГНОЗИРОВАНИЯ ОСТАТОЧНОГО РЕСУРСА СОСТАВНЫХ ЧАСТЕЙ АГРЕГАТОВ МАШИН // Москва, 1986 г. - 23 с
5. Уоссермен, Ф.Н. Нейрокомпьютерная техника Теория и практика [Электронный ресурс] // URL: http://cs.mipt.ru/docs/comp/rus/develop/neuro/neurokomp_technics/main.pdf (дата обращения 27.11.2015)
6. STATISTICA Electronic Manual