**АЗЕРБАЙДЖАНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСТИТЕТ НЕФТИ И ПРОМЫШЛЕННОСТИ**

**На правах рукописи**

Гусейнов Рустам Ровшан оглы

**“Разработка и исследование моделей прогнозирования основанных на технологи нейронных сетей для информационных систем электронного правительства”**

Специализация: 060632.1 – Информационные технологии и системная инженерия

**МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ**

**Научный руководитель: доц. Салимов В.Г.**

**БАКУ – 2024**

# **3.2 Решение практической задачи**

В качестве задачи, рассмотрим задачу прогнозирования роста регистрации автомобилей в 2024-2025 году на основании данных за прошлые периоды. Перед непосредственным решением этой задачи, необходимо подготовить таблицы записей для обучения, теста и прогнозирования. В качестве измерений будут выступать следующие параметры: год регистрации, регион регистрации, тип топлива и количество. Для моделирования работы сети, был подготовлен синтетический набор, сгенерированный с помощью самописного генератора случайных данных. Причина обращения к генератору заключается в отсутствии достоверных крупномасштабных данных в открытых источниках, что приводит к необходимости создания репрезентативного набора данных для целей настоящего исследования. Предоставим краткий обзор структуры генератора с подробным описанием ключевых параметров.

Начнем с описания модели(класса), характеризующего строку в наборе. Она состоит их следующих полей: год, регион, тип топлива, количество. Количество, при этом, ограничили 100 снизу и 3000 сверху (рисунок 3.4).

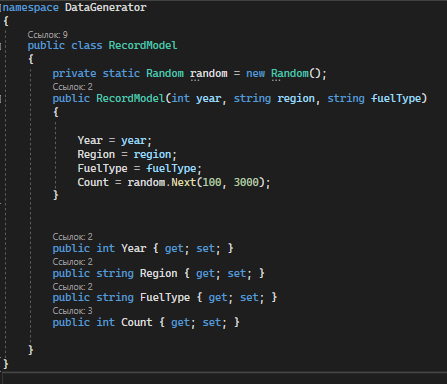


Рис.3.4 «Модель строки записи в наборе»

Теперь опишем устройство класса генератора. Как следует из рисунка 3.5, метод GenerateData() генератора создает по одной записи для каждой возможной комбинации типа год\регион\тип топлива, затем, на основании этих записей, создается и заполняется файл excel, посредством RecordsTable.

Рис.3.5 «Генератор записей набора»

Остается только класс самого набора данных RecordsTable (рисунок 3.6). Таблица представляется, как список элементов типа RecordModel. В классе присутствуют методы добавления записи в таблицу, отображения таблицы в окне консоли (использовалось для отладки) и метод экспорта таблицы в Excel

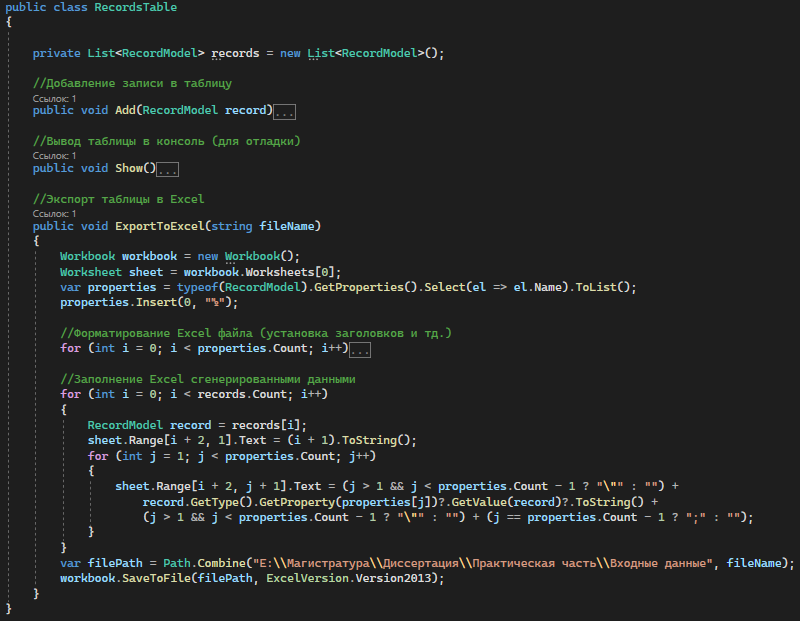


Рис.3.6 «Модель набора записей»

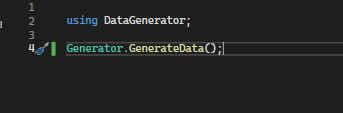


Рис.3.7 «Генерация Excel-я с данными»

В результате работы генератора получаем таблицу следующего вида

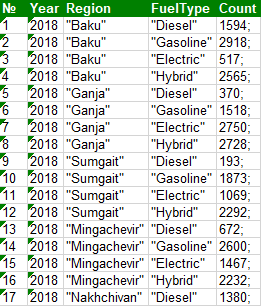
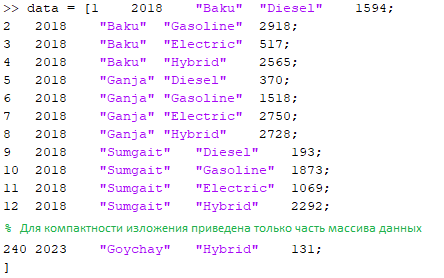
****

Рис.3.8 «Пример синтетического набора данных»

Генератор сгенерирует таблицу из 240 (6 лет \* 10 регионов \* 4 вида топлива) строк.

Перейдем непосредственно к построению нейронной сети для прогнозирования на основании сгенерированных данных. Будем пользоваться инструментом NNTool(Neural Network Tool) для Matlab.

Для начала перенесем данные из Excel в Matlab:



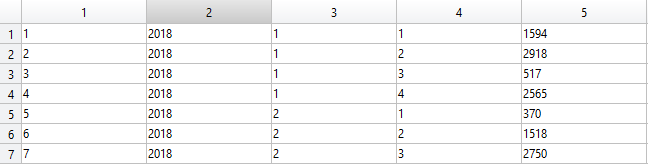
В списке переменных появилась data.

****

Преобразуем данные в вид, пригодный для использования нейросетью: требуется заменить записи в колонках «Регион» и «Тип топлива» на соответствующие уникальные числовые значения.



Теперь данные имеют следующий формат:



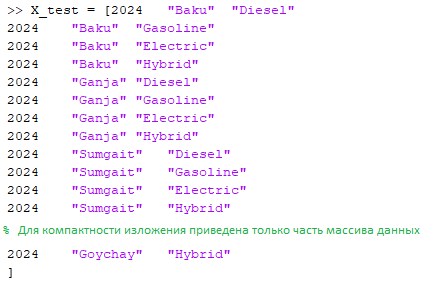
Извлекаем входные параметры (год, регион, тип топлива) и преобразуем строковый тип данных в численный



Извлекаем целевую переменную (количество) и преобразуем строковый тип данных в численный

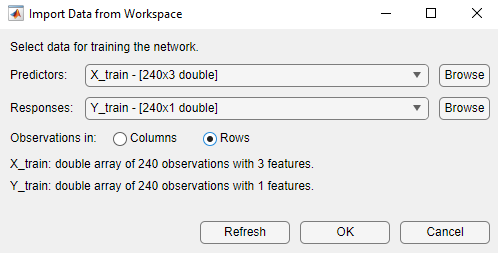


Создаем входные данные, на которых после будем тестировать уже обученную, к тому моменту, сеть.



С помощью X\_test = str2double(X\_test) приводим значения к численному виду.

Используя команду nftool вызываем UI для работы с инструментом Neural Network Fitting. Импортируем данные для тренировки сети: входные X\_train и выходные Y\_train:



Получаем сеть следующей конфигурации:

