МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение

высшего профессионального образования

«Ижевский государственный технический университет имени М.Т. Калашникова»

Факультет «Информатика и вычислительная техника»

Кафедра «Программное обеспечение»

Отчет по лабораторной работе №2

по дисциплине «Математические основы искусственного интеллекта»

на тему «Обзор программ моделирования искусственного интеллекта»

Weka

Выполнили:

студенты группы Б18-191-2 Гумметов Р.А.

Прилуков Н.А.

Принял: Коробейников А.В.

Ижевск 2021

1. Цель работы

Знакомство с программными пакетами моделирования искусственного интеллекта (ИИ).

1. Постановка задачи

Изучается основной подход к решению задач ИИ, реализованный в каждом выбранном программном пакете. Рассматривается один из примеров поставляемый с программой.

1. Теория

Weka

Weka — [свободное программное обеспечение](https://ru.wikipedia.org/wiki/%D0%A1%D0%B2%D0%BE%D0%B1%D0%BE%D0%B4%D0%BD%D0%BE%D0%B5_%D0%BF%D1%80%D0%BE%D0%B3%D1%80%D0%B0%D0%BC%D0%BC%D0%BD%D0%BE%D0%B5_%D0%BE%D0%B1%D0%B5%D1%81%D0%BF%D0%B5%D1%87%D0%B5%D0%BD%D0%B8%D0%B5) для анализа данных и машинного обучения, написанное на [Java](https://ru.wikipedia.org/wiki/Java) в [Университете Уаикато](https://ru.wikipedia.org/wiki/%D0%A3%D0%BD%D0%B8%D0%B2%D0%B5%D1%80%D1%81%D0%B8%D1%82%D0%B5%D1%82_%D0%A3%D0%B0%D0%B8%D0%BA%D0%B0%D1%82%D0%BE) ([Новая Зеландия](https://ru.wikipedia.org/wiki/%D0%9D%D0%BE%D0%B2%D0%B0%D1%8F_%D0%97%D0%B5%D0%BB%D0%B0%D0%BD%D0%B4%D0%B8%D1%8F)), распространяющееся по лицензии [GNU GPL](https://ru.wikipedia.org/wiki/GNU_GPL).

Weka представляет собой набор средств визуализации и алгоритмов для [интеллектуального анализа данных](https://ru.wikipedia.org/wiki/Data_mining) и решения [задач прогнозирования](https://ru.wikipedia.org/wiki/%D0%97%D0%B0%D0%B4%D0%B0%D1%87%D0%B8_%D0%BF%D1%80%D0%BE%D0%B3%D0%BD%D0%BE%D0%B7%D0%B8%D1%80%D0%BE%D0%B2%D0%B0%D0%BD%D0%B8%D1%8F), вместе с графической пользовательской оболочкой для доступа к ним.

Weka позволяет выполнять такие задачи анализа данных, как подготовку данных (preprocessing), [отбор признаков](https://ru.wikipedia.org/wiki/%D0%9E%D1%82%D0%B1%D0%BE%D1%80_%D0%BF%D1%80%D0%B8%D0%B7%D0%BD%D0%B0%D0%BA%D0%BE%D0%B2) ([англ.](https://ru.wikipedia.org/wiki/%D0%90%D0%BD%D0%B3%D0%BB%D0%B8%D0%B9%D1%81%D0%BA%D0%B8%D0%B9_%D1%8F%D0%B7%D1%8B%D0%BA) *feature selection*), [кластеризацию](https://ru.wikipedia.org/wiki/%D0%9A%D0%BB%D0%B0%D1%81%D1%82%D0%B5%D1%80%D0%B8%D0%B7%D0%B0%D1%86%D0%B8%D1%8F), [классификацию](https://ru.wikipedia.org/wiki/%D0%9A%D0%BB%D0%B0%D1%81%D1%81%D0%B8%D1%84%D0%B8%D0%BA%D0%B0%D1%86%D0%B8%D1%8F_(%D0%B7%D0%B0%D0%B4%D0%B0%D1%87%D0%B0)), [регрессионный анализ](https://ru.wikipedia.org/wiki/%D0%A0%D0%B5%D0%B3%D1%80%D0%B5%D1%81%D1%81%D0%B8%D0%BE%D0%BD%D0%BD%D1%8B%D0%B9_%D0%B0%D0%BD%D0%B0%D0%BB%D0%B8%D0%B7) и [визуализацию результатов](https://ru.wikipedia.org/wiki/%D0%92%D0%B8%D0%B7%D1%83%D0%B0%D0%BB%D0%B8%D0%B7%D0%B0%D1%86%D0%B8%D1%8F).

Weka содержит набор инструментов и алгоритмов визуализации для [анализа данных](https://en.wikipedia.org/wiki/Data_analysis) и [прогнозного моделирования](https://en.wikipedia.org/wiki/Predictive_modeling) , а также графические пользовательские интерфейсы для легкого доступа к этим функциям. Первоначальная версия Weka, отличная от Java, представляла собой интерфейс [Tcl](https://en.wikipedia.org/wiki/Tcl) / Tk для (в основном сторонних) алгоритмов моделирования, реализованных на других языках программирования, плюс утилиты [предварительной обработки данных](https://en.wikipedia.org/wiki/Data_preprocessing) на [C](https://en.wikipedia.org/wiki/C_(programming_language)) и основанная на [Makefile](https://en.wikipedia.org/wiki/Makefile) система для работающей машины.

Преимущества Weka:

* Бесплатная доступность по [Стандартной общественной лицензии GNU](https://en.wikipedia.org/wiki/GNU_General_Public_License) .
* Переносимость, поскольку он полностью реализован на [языке программирования Java](https://en.wikipedia.org/wiki/Java_programming_language) и, таким образом, работает практически на любой современной вычислительной платформе.
* Исчерпывающий набор методов предварительной обработки данных и моделирования.
* Простота использования благодаря графическому пользовательскому интерфейсу.

Weka поддерживает несколько стандартных задач [интеллектуального анализа данных](https://en.wikipedia.org/wiki/Data_mining) , а именно предварительную обработку данных, [кластеризацию](https://en.wikipedia.org/wiki/Data_clustering), [классификацию](https://en.wikipedia.org/wiki/Statistical_classification), [регрессию](https://en.wikipedia.org/wiki/Regression_analysis), визуализацию и [выбор функций](https://en.wikipedia.org/wiki/Feature_selection) . Все методы Weka основаны на предположении, что данные доступны в виде одного плоского файла или отношения, где каждая точка данных описывается фиксированным количеством атрибутов (обычно числовые или номинальные атрибуты, но также поддерживаются некоторые другие типы атрибутов) . Weka обеспечивает доступ к [базам данных](https://en.wikipedia.org/wiki/Database)[SQL](https://en.wikipedia.org/wiki/SQL) с помощью [Java Database Connectivity](https://en.wikipedia.org/wiki/Java_Database_Connectivity) и может обрабатывать результат, возвращаемый запросом к базе данных. Weka предоставляет доступ к [глубокому обучению](https://en.wikipedia.org/wiki/Deep_learning) с[Deeplearning4j](https://en.wikipedia.org/wiki/Deeplearning4j). Он не поддерживает интеллектуальный анализ данных, но есть отдельное программное обеспечение для преобразования набора связанных таблиц базы данных в единую таблицу, которая подходит для обработки с помощью Weka. Другой важной областью, которая в настоящее время не охвачена алгоритмами, включенными в дистрибутив Weka, является моделирование последовательностей.

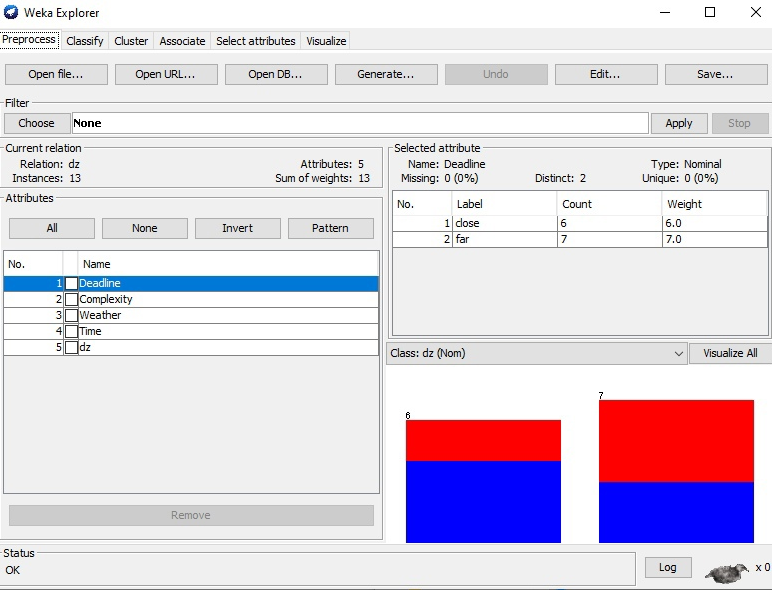


Рис. 1 - Weka

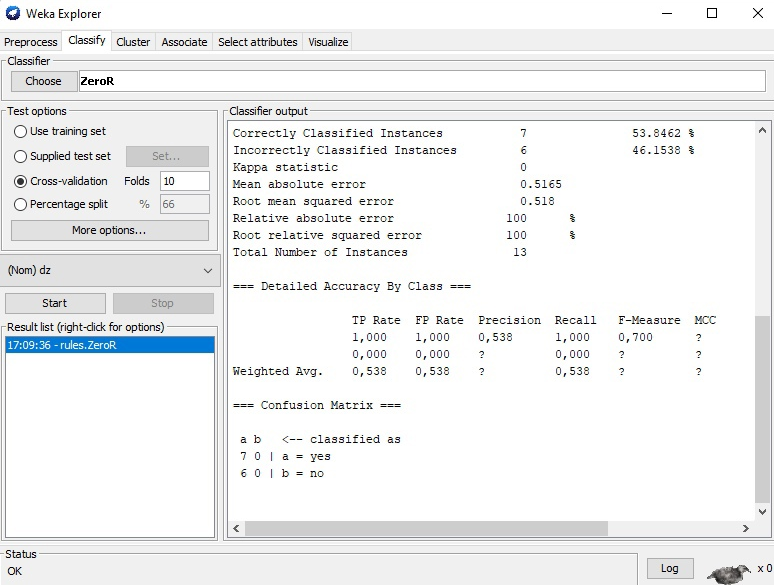


Рис. 2 - Weka

Яндекс.Автопоэт

Автопоэт Яндекса — робот, пишущий стихи на основе поисковых запросов пользователей. Руководитель службы аналитики поиска «Яндекса» Алексей Тихонов рассказал о создании «Автопоэта» — робота, который умеет автоматически составлять стихи в разных жанрах и из разного материала — из поисковых запросов и даже из заголовков новостей. Тихоново рассказал об алгоритме, который лежит в основе «Автопоэта», показал стихи, составленные по нестандартным материалам и написанные нейронной сетью в подражание известным поэтам, а также порассуждал о том, где проходит грань между творчеством человека и результатами компьютера.

Структура

Разберёмся со структурой — эту часть поэзии достаточно просто формализовать. Во-первых, мы ожидаем, что стих состоит из последовательности строк, сгруппированных в строфы. Во-вторых, для этих строк должны выполняться некоторые ритмические условия. В силлабическом, тоническом, силлабо-тоническом и других стихах эти условия формулируются по-разному, но так или иначе всё сводится к оценке ударности и длительности каждого слога. В-третьих, может потребоваться некоторая фонетическая согласованность между разными строками — это, например, может быть рифма или аллитерация.

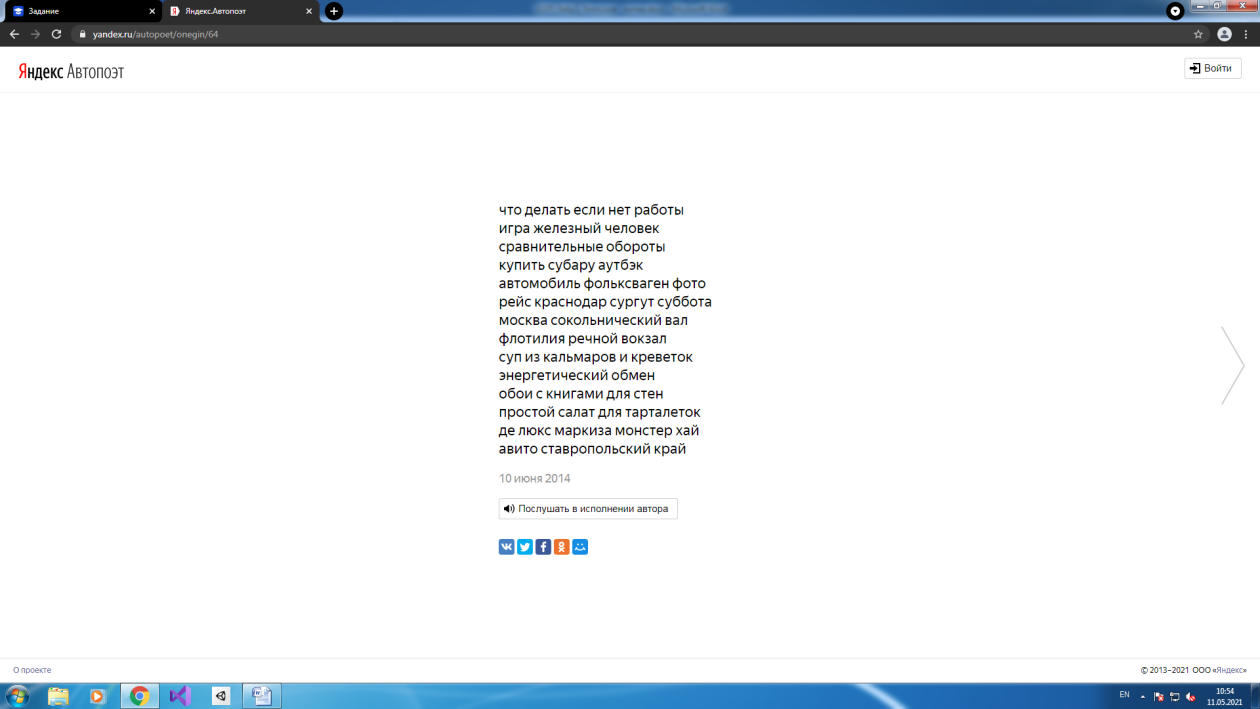


Рис. 3 - Яндекс.Автопоэт

Игра Life

Игра «Жизнь» — [клеточный автомат](https://ru.wikipedia.org/wiki/%D0%9A%D0%BB%D0%B5%D1%82%D0%BE%D1%87%D0%BD%D1%8B%D0%B9_%D0%B0%D0%B2%D1%82%D0%BE%D0%BC%D0%B0%D1%82), придуманный английским математиком [Джоном Конвеем](https://ru.wikipedia.org/wiki/%D0%9A%D0%BE%D0%BD%D0%B2%D0%B5%D0%B9,_%D0%94%D0%B6%D0%BE%D0%BD_%D0%A5%D0%BE%D1%80%D1%82%D0%BE%D0%BD) в [1970 году](https://ru.wikipedia.org/wiki/1970_%D0%B3%D0%BE%D0%B4).

Правила:

* Место действия этой игры — «вселенная» — это размеченная на клетки поверхность или плоскость — безграничная, ограниченная, или замкнутая (в пределе — бесконечная плоскость).
* Каждая клетка на этой поверхности может находиться в двух состояниях: быть «живой» (заполненной) или быть «мёртвой» (пустой). Клетка [имеет восемь соседей](https://ru.wikipedia.org/wiki/%D0%9E%D0%BA%D1%80%D0%B5%D1%81%D1%82%D0%BD%D0%BE%D1%81%D1%82%D1%8C_%D0%9C%D1%83%D1%80%D0%B0), окружающих её.
* Распределение живых клеток в начале игры называется первым поколением. Каждое следующее поколение рассчитывается на основе предыдущего по таким правилам:
  + в пустой (мёртвой) клетке, рядом с которой ровно три живые клетки, зарождается жизнь;
  + если у живой клетки есть две или три живые соседки, то эта клетка продолжает жить; в противном случае, если соседей меньше двух или больше трёх, клетка умирает («от одиночества» или «от перенаселённости»)
* Игра прекращается, если
  + на поле не останется ни одной «живой» клетки
  + конфигурация на очередном шаге в точности (без сдвигов и поворотов) повторит себя же на одном из более ранних шагов (складывается периодическая конфигурация)
  + при очередном шаге ни одна из клеток не меняет своего состояния (складывается стабильная конфигурация; предыдущее правило, вырожденное до одного шага назад)

Эти простые правила приводят к огромному разнообразию форм, которые могут возникнуть в игре.

Игрок не принимает прямого участия в игре, а лишь расставляет или генерирует начальную конфигурацию «живых» клеток, которые затем взаимодействуют согласно правилам уже без его участия (он является наблюдателем).

Компьютерная реализация

В компьютерных реализациях игры поле ограничено и (как правило) верхняя граница поля «соединена» с нижней, а левая граница — с правой, что представляет собой эмуляцию поверхности [тора](https://ru.wikipedia.org/wiki/%D0%A2%D0%BE%D1%80_(%D0%BF%D0%BE%D0%B2%D0%B5%D1%80%D1%85%D0%BD%D0%BE%D1%81%D1%82%D1%8C)), но на экране поле всегда отображается в виде равномерной сетки.

Простейший алгоритм «смены поколения» последовательно просматривает все ячейки решётки и для каждой ячейки подсчитывает соседей, определяя судьбу каждой клетки (не изменится, умрёт, родится). Такой простейший алгоритм использует два двумерных массива — один для текущего поколения, второй — для следующего.

Более сложный, но и более быстрый алгоритм составляет списки клеток для просмотра в последующем поколении; клетки, которые не могут измениться, в списки не вносятся. Например, если какая-либо клетка и ни одна из её соседей не поменялись на предыдущем ходу, то эта клетка не поменяется и на текущем ходу.

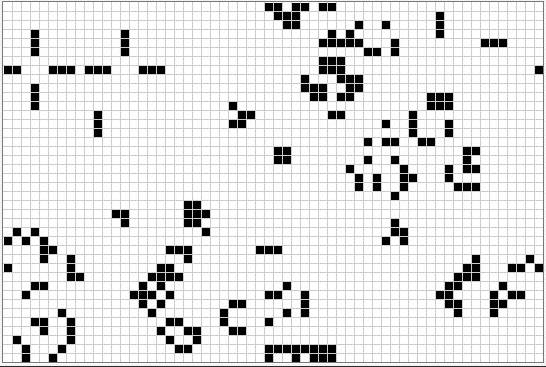


Рис. 4 - Игра Life

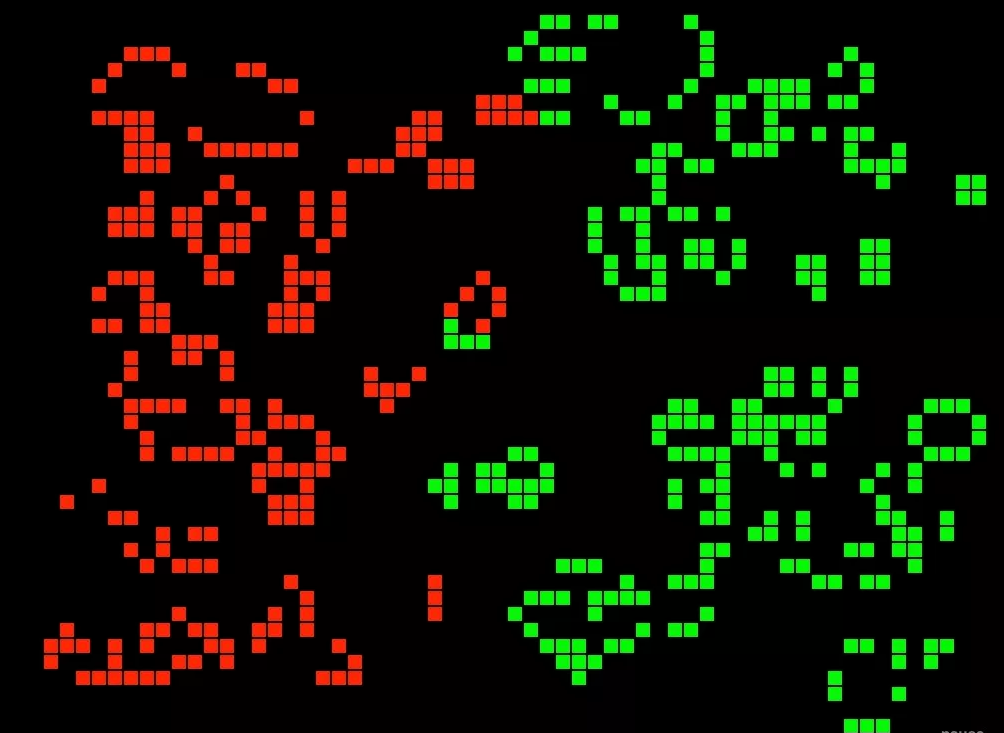


Рис. 5 - Игра Life