СОДРЖАНИЕ

ВВЕДЕНИЕ..............................................................................................................3

1 РАЗРАБОТКА АЛГОРИТМА РАСЧЕТА ПРОГНОЗА ПРОДАЖ (РПП)......5

1.1 Определение тренда................................................................................7

1.2 Логарифмическое представление тренда..............................................7

1.3 Линейное представление тренда..........................................................10

1.4 Полиномиальный тренд........................................................................15

1.5 Выбор наиболее точного расчета тренда............................................17

1.6 Алгоритм прогнозирования объема продаж пошагово.....................19

ЗАКЛЮЧЕНИЕ......................................................................................................21

**ВВЕДЕНИЕ**

Разрабатываемый продукт представляет собой экспертную систему ведения интернет торговли (ЭСИТ). ЭСИТ рассчитывает прогноз продаж путём использования уникального алгоритма, построенного в ходе разработки системы.

Алгоритм прогноза в ЭСИТ опирается на статистику продаж по прошедшим периодам. Для работы со статистическими данными был разработан модуль управления базой знаний (МУБЗ). Помимо внесения информации в базу данных и ее чтение, МУБЗ выполняет анализ актуальности данной информации. Информация является актуальной в течение определенного времени, обычно такое время составляет три прошедших периода. Понятие период для каждой группы товаров может быть разным, поэтому длительность одного периода выставляется индивидуально пользователем в настройках.

ЭСИТ так же предоставляет инструментарий, который автоматизирует управление бизнес процессами. Опираясь на результаты прогноза, система самостоятельно формирует список рекомендуемых товаров для закупки, то есть те товары, спрос на которые растет или, как минимум, остается стабильным, выводит списки пользователю на экран. При отсутствии такой автоматизации владельцу интернет-бизнеса придется самостоятельно выполнять этот рутинный анализ. Для интернет-магазинов с широким ассортиментом товаров ручной анализ обходится довольно дорого как по времени, так и по финансам.

Экспертная система ведения интернет-торговли - это не только запрограммированный алгоритм, а действительно целая система, состоящая из нескольких частей, каждая из которых выполняет свою функцию. На рисунке 1 изображена структура разрабатываемой системы.

Удаленный сервер

Источник   
данных

API

Клиентские системы

Алгоритм расчета

Рисунок 1 - Структура разрабатываемой системы

На сервере находится основная часть разработанного программного обеспечения, включая алгоритм прогнозирования, модуль управления базой знаний и непосредственно базу данных.

Источник данных представляет собой место хранения информации о товаре, это могут быть данные как со склада компании, занимающейся оптовой торговлей, так и данные с собственного склада.

API предназначен для того, чтобы наладить взаимодействие с интернет магазинами, которые решили подключить разработанную систему. Именно посредством API клиенты имеют возможность внедрять готовые функции от ЭСИТ в собственные системы, например вывод списка товара для интернет-магазина. Гибкость разработанного API позволяет внедрять функциональность не только в веб-версии своих проектов, а так же в мобильные или настольные приложения.

Клиентские системы представляют собой такие системы, в которые будет внедряться функциональность ЭСИТ. К ним относятся различного рода веб-сайты например интернет-магазины или порталы которые имют разделы с виртуальными витринами, а так же мобильные и настольные приложения, ориентированные на электронную коммерцию разработанные под различные операционные системы.

**1 РАЗРАБОТКА АЛГОРИТМА РАСЧЕТА ПРОГНОЗА ПРОДАЖ (РПП)**

На сегодняшний день разработано достаточно много различных технологий в области прогнозирования.

Активно применяются алгоритмы нейросетевого прогнозирования. Так же довольно часто применяются методы нечёткой логики. Большинство таких задач решается посредством применения методов исследования операций, к ним относятся: теория игр, имитационное моделирование, трендовый а так же регрессионный анализ. При разработке системы РПП был выбран один из возможных алгоритмов, строящих прогноз объёма продаж для товаров с сезонным характером.

Важно понимать, что такой величиной, как “сезон” в области прогнозирования могут быть любые временные колебания - сезонные вариации, например, в случае, когда проводится исследование товарооборота в течение одной недели, в качестве “сезона” выступает один рабочий день. К тому же, длительность цикла колебаний не всегда составляет один год - эта величина может варьироваться. После выявления величины цикла колебаний строится прогноз, используя для этого временной ряд, в котором была выявлена величина цикла с применением мультипликативных, а так же аддитивных моделей.

Аддитивная модель имеет вид:

|  |  |
| --- | --- |
| F = T + S ± E | (1) |

где:

*F* – прогнозируемое значение;

*Т* – тренд;

*S* – сезонная компонента;

*Е* – ошибка прогноза.

В формуле все значения указанны в денежных единицах.

Мультипликативные модели применяются в случае, когда во временных рядах наблюдается значение сезонной компоненты, которое является частью трендового значения. Такие модели имеют вид:

|  |  |
| --- | --- |
|  | (2) |

Ориентируясь на величину сезонной вариации возможно отличить аддитивную модель от мультипликативной. Для Аддитивной модели характерна практически постоянная сезонная вариация, в то время, как для мультипликативной модели сезонная вариация возрастает или убывает. На графиках это выражается в изменении амплитуды колебания сезонного фактора, пример изображен на рисунке 2.

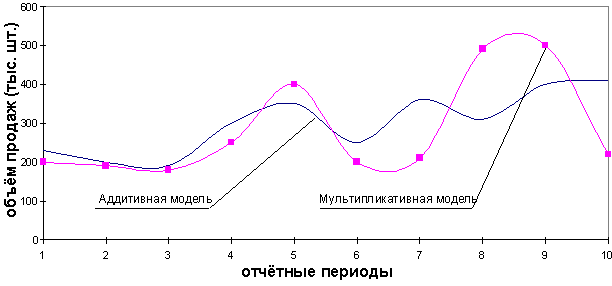


Рисунок 2 − Аддитивная и мультипликативная модели прогнозирования

**1.1 Определение тренда**

Существует множество различных взаимозаменяемых методов определения тренда. Наиболее часто встречающиеся и рекомендуемые в различных литературных источниках методы - это логарифмический, линейный, степенной, экспоненциальный и полиномиальный. Задачей является выбор метода, который дает наиболее точную оценку. Выбор метода в итоге, по результатам расчета, даст наиболее точный прогноз. Прежде, чем определиться с выбором метода, был изучен и опробован каждый из вышеперечисленных методов.

**1.2 Логарифмическое представление тренда**

• Логарифмический тренд имеет следующие свойства:

• когда b>0, уровни тренда растут с замедлением

• когда b<0, уровни тренда уменьшаются с замедлением

• абсолютные изменения уровней по модулю всегда уменьшаются со временем.

• цепные темпы изменения при t>? плавно приближаются к 100 % .

Напрашивается вывод о том, что логарифмический тренд, как и описанный ниже гиперболический тренд, отражает постепенно затухающий процесс изменений. Различие заключается в том, что затухание по гиперболе происходит довольно быстро при приближении к конечному пределу. В случае с логарифмическом трендом, затухающий процесс продолжается без ограничения значительно медленнее.

Система уравнений для определения коэффициентов уравнения регрессии имеет вид:

|  |  |
| --- | --- |
|  | (3) |

для функции вида:

|  |  |
| --- | --- |
|  | (4) |
|  | (5) |

При выполнении анализа рядов динамики большую роль играет выявление сезонных колебаний. Таким колебаниям свойственны практически постоянные изменения уровней ряда по внутригодовым периодам: месяцам, кварталам. Для того, чтобы выявить сезонные колебания, используют специальные показатели – индексы сезонности (Is).

Для ряда внутригодовой динамики, в которой основная тенденция роста незначительна (или она не наблюдается совсем), изучение сезонности базируется на методе постоянной средней, являющейся средней из всех рассматриваемых уровней. Один из способов получить индекс сезонности состоит в следующем: для каждого года отдельно рассчитывается средний уровень, после чего выполняется сопоставление (в процентах) уровня каждого месяца. Получаемое процентное отношение является индексом сезонности:

|  |  |
| --- | --- |
|  | (6) |

В большинстве случаев при выполнении моделирования рядов динамики с использованием полиномов или экспоненциальной функции не получают необходимых результатов, это происходит в силу того, что данные ряды динамики содержат в себе периодические колебания вокруг общей тенденции.

В таких случаях, как правило, приходится использовать гармонический анализ. Цель данного анализа заключается в нахождении и измерении периодических колебаний в рядах динамики. Функцию, которая была задана в каждой точке рассматриваемого интервала времени, представляют в виде бесконечного ряда синусоидальных и косинусоидальных функций.

Гармонический анализ - это операция выражения заданной периодической функции в виде ряда Фурье по гармоникам разных порядков. Каждый член ряда - это слагаемое постоянной величины с функциями синусов и косинусов некоторого периода.

Аппроксимация динамики финансовых явлений посредством ряда Фурье заключается в отборе таких гармонических колебаний, которые при наложении друг на друга отражают периодические колебания динамического ряда. Используя ряды Фурье, представляют динамику явлений в виде некоторой функции во времени, в которой слагаемые расставлены по убыванию периодов:

|  |  |
| --- | --- |
|  | (7) |

Методом наименьших квадратов рассчитывают параметры уравнений:

|  |  |
| --- | --- |
|  | (8) |
|  | (9) |
|  | (10) |

На графиках изображены варианты зависимостей факторного признака Y от факторного Х, где Х – временной фактор.

По данным таблицы 1, был построен график который изображен на рисунке 3.



Рисунок 3 - Логарифмический тренд

Выполнив расчет коэффициента детерминации, был получен результат.

**1.3 Линейное представление тренда**

Линейный тренд представлен в виде линейной зависимости анализируемой величины вида

|  |  |
| --- | --- |
|  | (11) |

где

y – исследуемая переменная (например, производительность);

x – значение, указывающее позицию (порядковый номер) года в периоде прогнозирования.

При выполнении линейной аппроксимации связи двух параметров, для нахождения эмпирических коэффициентов линейной функции чаще всего используют метод наименьших квадратов. Суть этого метода заключается в том, что линейная функция «наивысшего соответствия» проходит через точки графика, которые соответствуют минимуму суммы квадратов отклонений измеряемого параметра. Имеет вид:

|  |  |
| --- | --- |
|  | (12) |

Уравнения, рассчитывающие параметры линейной парной регрессии, имеют следующий вид:

|  |  |
| --- | --- |
|  | (13) |
|  | (14) |

где n – объем совокупности (исследуемой).

Константы a и b или значение коэффициента при переменной Х, а также свободного члена уравнения рассчитывается по следующей формуле:

|  |  |
| --- | --- |
|  | (15) |

а также:

|  |  |
| --- | --- |
|  | (16) |

В таблице 1 приведен пример вычисления линейного тренда по данным.

Таблица 1.

Вычисление линейного тренда

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Исходные данные | | XY |  |  |
| Период времени X | Объем продажи Y |
| 1 | 10 | 10 | 1 | 14,68 |
| 2 | 12 | 24 | 4 | 15,42 |
| 3 | 13 | 39 | 9 | 16,16 |
| 4 | 16 | 64 | 16 | 16,9 |
| 5 | 19 | 95 | 25 | 17,64 |
| 6 | 23 | 138 | 36 | 18,36 |
| 7 | 26 | 182 | 49 | 19,12 |
| 8 | 30 | 240 | 64 | 19,86 |
| 9 | 28 | 252 | 81 | 20,6 |
| 10 | 18 | 180 | 100 | 21,34 |
| 11 | 16 | 176 | 121 | 22,08 |
| 12 | 14 | 168 | 144 | 22,82 |
|  |  |  |  |  |
| Среднее 6,5 | 18,75 |  |  |  |

, ; , ; ;

= =;

= 18,75 − 0,74 \* 6,5 = 18,25 − 4,81 = 13,94

уравнение тренда:

|  |  |
| --- | --- |
| = 13,94 + 0,74 | (17) |

Методы сглаживания колебаний. В случае, когда наблюдаются большие расхождения между соседними значениями, тренд, рассчитанный методом регрессии, сложно анализировать. В случае, когда при построении прогноза ряд состоит из данных с высокой амплитудой колебаний между ближайшими значениями, выполняется сглаживание с использованием специальных методов. Таких методов существует довольно много, самые известные среди них: метод экспоненциального сглаживания и метод скользящих средних.

Метод скользящих средних или МСС. МСС сглаживает ряд значений, что позволяет выявить тренд. Для проведения расчетов посредством данного метода используется среднее арифметическое значение определенного количества значений, например трехточечное скользящее среднее. Для этого берутся первые три значения из списка, например данные за январь, февраль и март, затем считается среднее арифметическое значение, которое ставится в центре диапазона, то есть в феврале. Далее выполняется сдвиг на один месяц и берутся три следующих значения, то есть февраль, март, апрель, затем считается среднее арифметическое значение уже между данными тремя значениями. По такому принципу обрабатывается весь ряд. Новые значения, которые были получены в результате полной обработки ряда, составляют новый ряд. Чем больше ряд содержит значений, тем сильнее выполняется сглаживание. Пример построения тренда представлен в таблице 2.

Таблица 2

Расчет тренда методом трехточечного скользящего среднего

|  |  |  |
| --- | --- | --- |
| Месяц | Продажи, тыс. ед. | Расчет |
| Янв. | 11 | (11 + 13 + 14) / 3 = 12,67 |
| Февр. | 13 | (13 + 14 + 17) / 3 = 14,67 |
| Месяц | Продажи, тыс. ед. | Расчет |
| Март | 14 | (14 + 17 + 20) / 3 = 17 |
| Апр. | 17 | (17 + 20 + 24) / 3 = 20,33 |
| Май | 20 | (20 + 24 + 27) / 3 = 23,67 |

Продолжение таблицы 2

Расчет тренда методом трехточечного скользящего среднего

|  |  |  |
| --- | --- | --- |
| Месяц | Продажи, тыс. ед. | Расчет |
| Июнь | 24 | (24 + 27 + 31) / 3 = 27,33 |
| Июль | 27 | (24 + 27 + 31) / 3 = 27,33 |
| Авг. | 31 | (27 + 31 + 29) / 3 = 29 |
| Сент. | 29 | (31 + 29 + 19) / 3 = 26,33 |
| Окт. | 19 | (29 + 19 + 17) / 3 = 22,67 |
| Нояб. | 17 | (19 + 17 + 15) / 3 = 17 |
| Дек. | 15 |  |

Динамика колебаний исходных значений, которые были получены методом скользящего среднего, изображена на рисунке #. Результаты сравнения графиков рядов исходных значений и трехточечных скользящих средних показывают наличие возможности сглаживания колебания. Чем больше точек содержится в диапазоне вычисления, тем отчетливее будет вырисовываться тренд. Стоит отметить тот факт, что операция укрупнения диапазона ведет к уменьшению количества конечных значений, что, в свою очередь, приводит к менее точному прогнозу.



Рисунок 4 − динамика изменения объема продаж

На рисунке 4 изображена динамика изменения объема продаж:

ряд №3 - исходные данные;

ряд №4 - скользящее среднее;

ряд №2 - экспоненциальное среднее;

ряд №1 - тренд, построенный методом регрессии.

Метод экспоненциального сглаживания. Еще один подход к уменьшению разброса значений ряда. Такое значение ему было присвоено в силу того, что каждое значение периодов, уходящих в прошлое, уменьшается на определенный множитель, то есть 1 – α. Для расчета каждого значения используется формуа:

|  |  |
| --- | --- |
| St =aYt +(1−α)St−1 | (19) |

где:

St – текущее сглаженное значение;

Yt – текущее значение временного ряда;

St – 1 – предыдущее сглаженное значение;

α – сглаживающая константа, 0 ≤ α ≤ 1.

Чем меньше значение константы α, тем менее оно чувствительно к изменениям тренда в данном временном ряду.

Посчитав коэффициент детерминации, был получен результат

= 0,016.

**1.4 Полиномиальный тренд**

**Полиномиальный тренд** применяется для описания значений временных рядов, попеременно возрастающих и убывающих. Полином отлично подходит для анализа большого набора данных нестабильной величины чаще всего при продажи сезонных товаров. Полиномиальный трендописывает данные, плавно изменяющиеся в разных направлениях. При использовании полиномиального тренда пользователю всегда необходимо задать порядок полинома.

Полином это степенная функция, в данном случае может иметь от двух до шести порядков, из-за чего уравнение полиномиального тренда имеет разный вид. Так уравнение полинома второго порядка имеет вид:

|  |  |
| --- | --- |
|  | (20) |

Уравнение полинома третьего порядка выглядит:

|  |  |
| --- | --- |
|  | (21) |

Уравнение для полинома четвертого порядка:

|  |  |
| --- | --- |
|  | (22) |

Уравнение для полинома пятого порядка:

|  |  |
| --- | --- |
|  | (23) |

Уравнение для полинома шестого порядка:

|  |  |
| --- | --- |
|  | (24) |

Степень полинома определяет количество экстремумов или пиков, то есть число максимальных и минимальных значений на анализируемом промежутке времени.

Например у полинома второй степени может быть только один экстремум.

На рисунке 5 изображен график полиномиального представления тренда, на котором наблюдается только один экстремум.

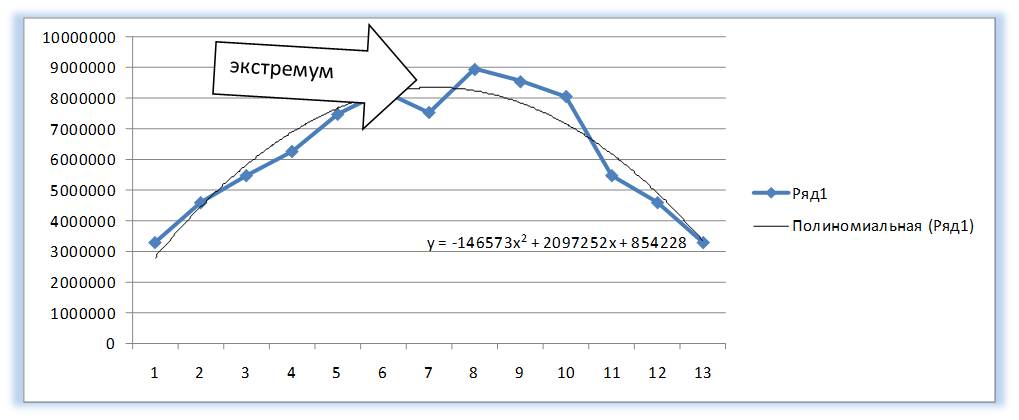


Рисунок 5 Полином второй степени

У полинома третьей степени может быть от одного до двух экстремумов.

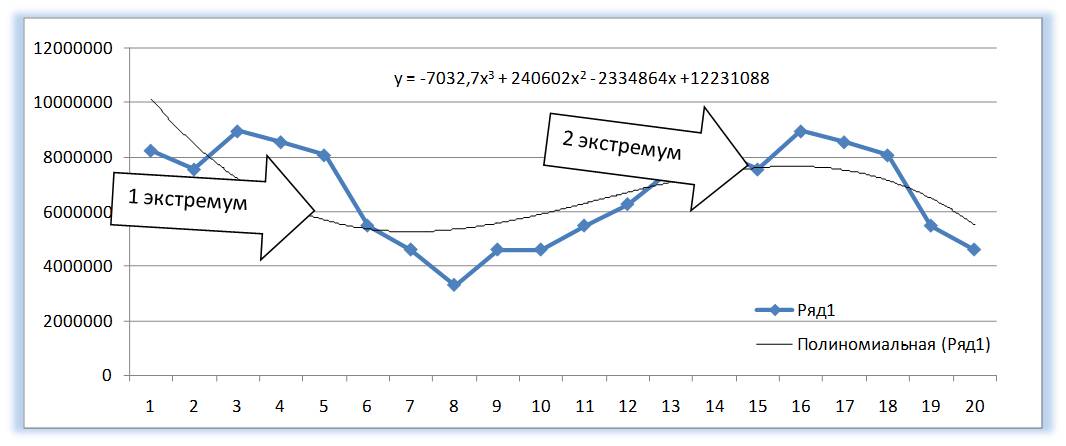


Рисунок 6 Полином третей степени

Этот же принцип аналогичен и для остальных полиномов, так например полином шестой степени может иметь до пяти экстремумов.

**1.5 Выбор наиболее точного расчета тренда**

Согласно результатам исследований и проведенных экспериментов был выбран полиномиальный тренд. Результаты показали, что данный тренд в сравнении с таким трендами, как степенной, логарифмический или экспоненциальный, оказался наиболее эффективным, так как он наиболее удовлетворительно аппроксимирует фактические значения и имеет самый высокий коэффициент детерминации:

• логарифмический R2 = 0,0166;

• степенной R2 =0,0197;

• экспоненциальный R2 =8Е-05.

Полиномиальный же R2 = 0,7435, что позволяет в разы уменьшить вероятность ошибки при прогнозировании.

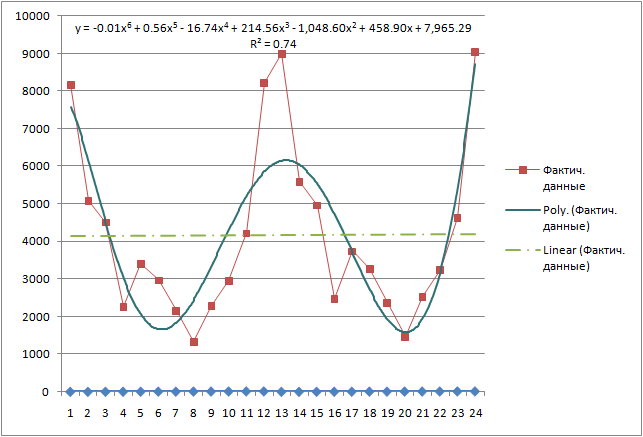
****

Рисунок 7 Сравнительный анализ полиномиального и линейного тренда

На рисунке 7 изображена диаграмма, изучив которую можно сделать вывод о том, что полиномиальный тренд выполняет аппроксимацию фактических данных гораздо эффективнее, нежели другие часто используемые тренды, например линейный.

**1.6 Алгоритм прогнозирования объема продаж пошагово**

Приведенный ниже алгоритм прогнозирования объема продаж имеет сезонный характер и состоит из нескольких шагов.

Первым шагом является определение тренда. Существует множество способов расчета тренда, самые распространенные из них описаны выше. С ними был проведен эксперимент, по результатам которого выбран полиномиальный тренд, так как именно он наилучшим образом позволяет сократить ошибку прогнозной модели.

Следующим шагом выполняется вычитание значения тренда из фактических значений объемов продаж. Это позволяет определить величину сезонной компоненты. После того, как сезонная компонента определена, ее необходимо откорректировать таким образом, чтобы сумма фактических значений и сезонной компоненты была равна нулю.

После определения сезонной компоненты и ее корректировки выполняется расчет ошибки модели. Для этого нужно посчитать разницу между фактическими значениями и значениями модели.

Когда посчитаны все значения, выполняется построение модели прогнозирования по формуле:

|  |  |
| --- | --- |
| F = T + S ± E | (25) |

где:

F– прогнозируемое значение;

Т– тренд;

S – сезонная компонента;

Е - ошибка модели.

На основе построенной прогнозной модели рассчитывается окончательный прогноз продаж. Для сглаживания ряда был выбран метод экспоненциального сглаживания, так как он позволяет учесть весьма вероятные изменения экономических тенденций в будущем.

|  |  |
| --- | --- |
|  | (26) |

**где:**

– прогнозное значение объёма продаж;

– фактическое значение объёма продаж в предыдущем году;

– значение модели;

**–** константа сглаживания.

Результаты применения метода на практике показали ряд его особенностей:

• Для построения прогноза достаточно указать величину сезона. Исследования показывают, что у большинства продуктов присутствует именно сезонный характер.

Важно понимать то, что величина сезона для каждого вида товара индивидуальна. Для одной группы товаров это может быть неделя, для другой группы - год.

• Использование полиномиального тренда позволяет значительно сократить ошибку по сравнению с другими трендами.

• При условии внесения достаточного количества данных (в зависимости от величины сезона), метод выполняет эффективную аппроксимацию и может быть полезен в построении стратегии ведения продаж.

**ЗАКЛЮЧЕНИЕ**

При проектировании экспертной системы ведения торговли (ЭСИТ) был исследован ряд математических методов прогнозирования продаж на основе трендового анализа. По результатам экспериментов, проведенных с этими методами был определен самый эффективный. который был взят за основу при разработки ЭСИТ.

При исследовании рассматривались такие представления тренда как: логарифмическое, линейное, полиномиальное. Результаты исследования показали, что использование именно полиномиального представления тренда дает наиболее высокий эффект.

Был испытан спроектированный алгоритм, а так же на основе реальных данных по продажам был построен прогноз. Результаты сравнения прогноза и реальных продаж за период, на который был построен прогноз, показали, что алгоритм строит довольно точный прогноз. В получаемом прогнозе выявлены ошибочные отклонения, однако данные отклонения превышают допустимую норму, а значит, использования данного алгоритма для прогнозирования оправдано.

Спроектирована непосредственна сама ЭСИТ, составлены блок-схемы, описывающие работу всей системы и показывающие применение на практике разработанного алгоритма прогнозирования, который был внедрен в разрабатываемую систему.

В целом результаты проведенной работы показали, что заявленные технические требования к ЭСИТ реализуемы. Доступные современны технологии позволяют решить поставленные задачи при разработке системы, а проведенные экономические расчеты подтвердили целесообразность разработки экспертной системы ведения торговли.