

Министерство образования и науки Российской Федерации
федеральное государственное автономное образовательное учреждение
высшего образования
**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**

Школа: Инженерная школа информационных технологий и робототехники
Направление подготовки: Информационные системы и технологии
Отделение школы (НОЦ): Отделение информационных технологий

БАКАЛАВРСКАЯ РАБОТА

Тема работы
ПРОГНОЗИРОВАНИЕ ВРЕМЕННЫХ РЯДОВ НА ОСНОВЕ ШТРАФНОГО Р-СПЛАЙНА

УДК 519.246-047.72:519.68

Студент

Группа	ФИО	Подпись	Дата
8И4А	Рожкова Надежда Евгеньевна		

Руководитель

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОИТ	Кочегурова Е.А.	К.Т.Н.		

КОНСУЛЬТАНТЫ:

По разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Ст. преподаватель ОСГН	Хаперская А.В.			

По разделу «Социальная ответственность»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОКД	Штейнле А.В.	К.М.Н.		

ДОПУСТИТЬ К ЗАЩИТЕ:

Руководитель ООП	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОИТ	Цапко И.В.	К.Т.Н.		

Томск – 2018 г.

РЕЗУЛЬТАТЫ ОБУЧЕНИЯ

Код результатов	Результат обучения (выпускник должен быть готов)
Профессиональные и общепрофессиональные компетенции	
P1	Применять базовые и специальные естественнонаучные и математические знания для комплексной инженерной деятельности по созданию, внедрению и эксплуатации геоинформационных систем и технологий, а также информационных систем и технологий в бизнесе.
P2	Применять базовые и специальные знания в области современных информационных технологий для решения инженерных задач.
P3	Ставить и решать задачи комплексного анализа, связанные с созданием геоинформационных систем и технологий, информационных систем в бизнесе, с использованием базовых и специальных знаний, современных аналитических методов и моделей.
P4	Выполнять комплексные инженерные проекты по созданию информационных систем и технологий, а также средств их реализации (информационных, методических, математических, алгоритмических, технических и программных).
P5	Проводить теоретические и экспериментальные исследования, включающие поиск и изучение необходимой научно-технической информации, математическое моделирование, проведение эксперимента, анализ и интерпретация полученных данных, в области создания геоинформационных систем и технологий, а также информационных систем и технологий в бизнесе.
P6	Внедрять, эксплуатировать и обслуживать современные геоинформационные системы и технологии, информационные системы и технологии в бизнесе, обеспечивать их высокую эффективность, соблюдать правила охраны здоровья, безопасность труда, выполнять требования по защите окружающей среды.
Универсальные (общекультурные) компетенции	
P7	Использовать базовые и специальные знания в области проектного менеджмента для ведения комплексной инженерной деятельности.
P8	Осуществлять коммуникации в профессиональной среде и в обществе в целом. Владеть иностранным языком (углублённый английский язык), позволяющем работать в иноязычной среде, разрабатывать документацию, презентовать и защищать результаты комплексной инженерной деятельности.
P8	Эффективно работать индивидуально и в качестве члена команды, состоящей из специалистов различных направлений и квалификаций.
P10	Демонстрировать личную ответственность за результаты работы и готовность следовать профессиональной этике и нормам ведения комплексной инженерной деятельности
P11	Демонстрировать знания правовых, социальных, экологических и культурных аспектов комплексной инженерной деятельности, а также готовность к достижению должного уровня физической подготовленности для обеспечения полноценной социальной и профессиональной деятельности.

Министерство образования и науки Российской Федерации
 федеральное государственное автономное образовательное учреждение
 высшего образования
**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
 ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**

Школа: Инженерная школа информационных технологий и робототехники
 Направление подготовки: Информационные системы и технологии
 Отделение школы (НОЦ): Отделение информационных технологий

УТВЕРЖДАЮ:
 Руководитель ООП
 _____ Кочегурова Е.А.
 (Подпись) (Дата) (Ф.И.О.)

ЗАДАНИЕ
на выполнение выпускной квалификационной работы

В форме:

бакалаврской работы

(бакалаврской работы, дипломного проекта/работы, магистерской диссертации)

Студенту:

Группа	ФИО
8И4А	Рожковой Надежде Евгеньевне

Тема работы:

Прогнозирование временных рядов на основе штрафного Р-сплайна	
Утверждена приказом директора (дата, номер)	№2063/с от 23.03.2018

Срок сдачи студентом выполненной работы:	04.06.2018
--	------------

ТЕХНИЧЕСКОЕ ЗАДАНИЕ:

Исходные данные к работе	Реализация модели прогнозирования на основе штрафных Р-сплайнов и обеспечение качества прогнозирования с помощью метода оптимизации.
Перечень подлежащих исследованию, проектированию и разработке вопросов	Обзор существующих методов прогнозирования, выбор метода оптимизации, анализ оценок качества прогнозирования. Исследование и анализ результатов прогноза, полученного с использованием, основанной на штрафных Р-сплайнах.
Перечень графического материала	Презентация в формате *.pptx
Консультанты по разделам выпускной квалификационной работы <i>(с указанием разделов)</i>	
Раздел	Консультант

Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	Хаперская А.В.
Социальная ответственность	Штейнле А.В.
Названия разделов, которые должны быть написаны на русском и иностранном языках:	
Заключение	

Дата выдачи задания на выполнение выпускной квалификационной работы по линейному графику	01.03.2018
---	------------

Задание выдал руководитель:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОИТ	Кочегурова Е.А.	к.т.н.		

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8И4А	Рожкова Надежда Евгеньевна		

Министерство образования и науки Российской Федерации
федеральное государственное автономное образовательное учреждение
высшего образования
**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**

Инженерная школа Информационных технологий и робототехники
Направление подготовки (специальность) Информационные системы и технологии
Уровень образования Бакалавриат
Отделение школы (НОЦ) Информационных технологий
Период выполнения весенний семестр 2017/2018 учебного года

Форма представления работы:

Бакалаврская работа
(бакалаврская работа, дипломный проект/работа, магистерская диссертация)

КАЛЕНДАРНЫЙ РЕЙТИНГ-ПЛАН
выполнения выпускной квалификационной работы

Студент

Срок сдачи студентом выполненной работы:	04.06.2018
--	------------

Дата контроля	Название раздела (модуля) / вид работы (исследования)	Максимальный балл раздела (модуля)
31.05.2018	Основная часть	75
21.05.2018	Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	15
19.05.2018	Социальная ответственность	10

Составил преподаватель:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОИТ	Кочегурова Е.А.	к.т.н.		

СОГЛАСОВАНО:

Руководитель ООП	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОИТ	Цапко И.В.	к.т.н.		

Томск – 2018 г.

**ЗАДАНИЕ ДЛЯ РАЗДЕЛА
«ФИНАНСОВЫЙ МЕНЕДЖМЕНТ, РЕСУРСООБЪЕКТИВНОСТЬ И
РЕСУРСОСБЕРЕЖЕНИЕ»**

Студентке:

Группа	ФИО
8И4А	Рожкова Надежда Евгеньевна

Школа	ИШИТР	Отделение	Информационных технологий
Уровень образования	Бакалавриат	Направление/специальность	Информационные системы и технологии

Исходные данные к разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»:

1. <i>Стоимость ресурсов научного исследования (НИ): материально-технических, энергетических, финансовых, информационных и человеческих</i>	Работа с информацией, представленной в российских и иностранных научных публикациях, аналитических материалах, статистических бюллетенях и изданиях, нормативно-правовых документах.
2. <i>Нормы и нормативы расходования ресурсов</i>	
3. <i>Используемая система налогообложения, ставки налогов, отчислений, дисконтирования и кредитования</i>	

Перечень вопросов, подлежащих исследованию, проектированию и разработке:

1. <i>Оценка коммерческого потенциала, перспективности и альтернатив проведения НИ с позиции ресурсоэффективности и ресурсосбережения</i>	Оценка потенциальных потребителей, анализ конкурентных решений, QuaD и SWOT-анализ научного исследования, определение возможных альтернатив с использованием морфологического подхода.
2. <i>Планирование и формирование бюджета научных исследований</i>	Планирование этапов работ, определение трудоемкости работы, построение календарного графика (диаграмма Ганта), расчет бюджета научно-технического проекта
3. <i>Определение ресурсной (ресурсосберегающей), финансовой, бюджетной, социальной и экономической эффективности исследования</i>	Оценка ресурсной, финансовой и экономической эффективности научно-исследовательского проекта.

Перечень графического материала (с точным указанием обязательных чертежей):

1. <i>Оценка конкурентоспособности технических решений</i>
2. <i>Матрица SWOT</i>
3. <i>Альтернативы проведения НИ</i>
4. <i>График проведения и бюджет НИ</i>
5. <i>Оценка ресурсной, финансовой и экономической эффективности НИ</i>

Дата выдачи задания для раздела по линейному графику	22.01.2018
---	------------

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Старший преподаватель ОСГН	Хаперская А. В.			

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8И4А	Рожкова Надежда Евгеньевна		

ЗАДАНИЕ ДЛЯ РАЗДЕЛА «СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ»

Студенту:

Группа	ФИО
8И4А	Рожкова Надежда Евгеньевна

Школа	ИШИТР	Отделение	информационных технологий
Уровень образования	Бакалавриат	Направление/специальность	Информационные системы и технологии

Исходные данные к разделу «Социальная ответственность»:

1. Характеристика объекта исследования (вещество, материал, прибор, алгоритм, методика, рабочая зона) и области его применения	1. Разработка программного приложения для краткосрочного прогнозирования временных рядов на основе штрафного Р-сплайна. Приложение предназначено для применения на компьютерах с ОС Windows.
--	--

Перечень вопросов, подлежащих исследованию, проектированию и разработке:

1. Производственная безопасность 1.1. Анализ выявленных вредных факторов при разработке и эксплуатации проектируемого решения. 1.2. Анализ выявленных опасных факторов при разработке и эксплуатации проектируемого решения.	Выявление и анализ следующих вредных факторов: отклонение показателей микроклимата; недостаточная освещенность рабочей зоны. Выявление и анализ следующих опасных факторов: электрический ток, пожаровзрывобезопасность
2. Экологическая безопасность: 2.1. Анализ воздействия объекта на атмосферу (выбросы), гидросферу (сбросы), литосферу (отходы); 2.2. Разработать решения по обеспечению экологической безопасности со ссылками на НТД по охране окружающей среды.	Выявление и анализ негативно влияющих на окружающую среду факторов при эксплуатации персонального компьютера. Определение решения по обеспечению экологической безопасности согласно нормативным документам.
3. Безопасность в чрезвычайных ситуациях: 3.1. Перечень возможных ЧС при разработке и эксплуатации проектируемого решения; 3.2. Выбор наиболее типичной ЧС; 3.3. Разработка превентивных мер по предупреждению ЧС; 3.4. Разработка действий в результате возникшей ЧС и мер по ликвидации её последствий.	Перечень возможных чрезвычайных ситуаций, которые могут возникнуть при работе в офисном помещении. Способы защиты от пожара и ликвидация последствий
4. Правовые и организационные вопросы обеспечения безопасности: 4.1. Специальные правовые нормы трудового законодательства; 4.2. Организационные мероприятия при компоновке рабочей зоны.	Организационные мероприятия по обеспечению безопасности работников, трудящихся за персональным компьютером, правовые нормы трудового законодательства Организационные мероприятия при компоновке рабочей зоны.

Дата выдачи задания для раздела по линейному графику	22.01.2018
---	------------

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОКД	Штейнле А.В.	К.М.Н.		

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8И4А	Рожкова Надежда Евгеньевна		

РЕФЕРАТ

Выпускная квалификационная работа содержит 105 страниц, 21 рисунок, 26 таблиц, 41 источник, 5 приложений.

Ключевые слова: краткосрочное прогнозирование, штрафной сплайн, р-сплайн, модель прогнозирования, оптимизация.

Объектом исследования является математическая модель краткосрочного прогнозирования, основанная на штрафных Р-сплайнах.

Цель работы – реализовать модель прогнозирования, основанную на штрафных Р-сплайнах в виде программного приложения и обеспечить высокую оценку качества прогнозирования, определяя оптимальные входные параметры сплайна.

В процессе исследования был проведен обзор современных методов прогнозирования, проанализированы методы оптимизации и выбран наиболее подходящий для решения задачи определения оптимальных параметров, рассмотрены методы оценок качества прогнозирования.

В результате исследования было спроектировано приложение, реализующее модель прогнозирования на основе штрафных Р-сплайнов и обеспечивающее определение оптимальных параметров с применением генетического алгоритма, а также был произведен анализ результатов прогнозирования.

Область применения: исследовательская научная и образовательная деятельность.

В рамках развития проекта в будущем планируется следующее:

- Расширение функциональных возможностей
- Оптимизация исходного кода
- Дальнейшее исследование и поиск областей применения данной модели прогнозирования.

ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ

WPF – Windows Presentation Foundation

XAML – eXtensible Application Markup Language

ARIMA – Autoregressive Integrated Moving Average

RNN – Recurrent Neural Network

LSTM – Long/Short Term Memory

GRU – Gated Recurrent Unit

MSPE – Mean Squared Percentage Error

MSE – Mean Squared Error

MAE – Mean Absolute Error

LINQ – Language Integrated Query Language

МНК – Метод Наименьших Квадратов

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ.....	13
1 Анализ предметной области	14
1.1 Прикладные области краткосрочного прогнозирования	14
1.2 Обзор методов краткосрочного прогнозирования	16
1.2.1 Основные статистические методы прогнозирования	17
1.2.2 Использование нейронных сетей в задаче прогнозирования	22
1.2.3 Использование сплайн-функций в задаче прогнозирования.	24
2 Разработка метода прогнозирования на основе штрафного P-сплайна	26
2.1 Расчетная схема и режимы функционирования штрафного P- сплайна	26
2.2 Показатели качества прогнозирования.....	30
3 Оптимизация параметров сплайна	35
3.1 Сравнение методов оптимизации параметров сплайна	36
3.2 Использование генетического алгоритма для задачи оптимизации	38
4 Разработка программного приложения	43
4.1 Среда разработки и язык программирования	43
4.2 Структура программного приложения	44
4.3 Функциональные требования к программному приложению.....	46
4.4 Интерфейс пользователя	47
5 Исследование работы программного приложения.....	49
5.1 Анализ результатов краткосрочного прогнозирования на основе штрафных P-сплайнов	49
5.2 Анализ результатов качества прогнозирования с использованием оптимизационного алгоритма.....	53

6. Финансовый менеджмент, ресурсоэффективность и ресурсосбережение.....	58
6.1. Оценка коммерческого потенциала, перспективности и альтернатив проведения НИ с позиции ресурсоэффективности и ресурсосбережения	59
6.1.1. Потенциальные потребители результатов исследования.....	59
6.1.2. Анализ конкурентных технических решений.....	60
6.1.3. Технология QuaD.....	62
6.1.4. SWOT-анализ.....	63
6.1.5. Определение возможных альтернатив проведения научных исследований	65
6.2. Планирование и формирование бюджета научных исследований	66
6.2.1. Структура работ в рамках научного исследования.....	66
6.2.2. Определение трудоемкости выполнения работ.....	66
6.2.3. Разработка графика проведения научного исследования.....	68
6.2.4. Бюджет научно-технического исследования	69
6.3. Определение ресурсной (ресурсосберегающей), финансовой, бюджетной, социальной и экономической эффективности исследования .	73
7. Социальная ответственность	76
7.1. Производственная безопасность	76
7.1.1. Вредные опасные факторы	77
7.1.2. Опасные производственные факторы.....	81
7.2. Экологическая безопасность	83
7.2.1. Анализ воздействия продукта на окружающую среду	83
7.2.2. Решения по обеспечению экологической безопасности.....	84

7.3. Безопасность в чрезвычайных ситуациях	84
7.3.1. Перечень возможных ЧС при разработке и эксплуатации научно-исследовательского проекта	84
7.3.2. Разработка действий в результате возникшей ЧС и меры по ликвидации ее последствий	86
7.4. Правовые и организационные вопросы обеспечения безопасности	88
7.4.1. Специальные правовые нормы трудового законодательства	88
7.4.2. Организационные мероприятия при компоновке рабочей зоны	89
ЗАКЛЮЧЕНИЕ	91
CONCLUSION	92
Список использованных источников	93
ПРИЛОЖЕНИЕ А	98
ПРИЛОЖЕНИЕ Б	99
ПРИЛОЖЕНИЕ В	100
ПРИЛОЖЕНИЕ Г	101
ПРИЛОЖЕНИЕ Д	103

ВВЕДЕНИЕ

На данный момент существует множество методов краткосрочного прогнозирования временных рядов, применяемых в различных областях, начиная с экономики и финансов и заканчивая промышленностью и электроэнергетикой. Данные методы и модели прогнозирования имеют свои достоинства и недостатки, но, прежде всего, ориентированы на повышение точности предсказаний. При этом повышение качества прогнозирования негативно влияет на такие факторы прогнозирования, как время прогнозирования и необходимость предварительной работы с входными данными, что затрудняет прогнозирование в масштабе реального времени.

Однако, наибольшую ценность в ходе наблюдения за объектом несут данные, получаемые в режиме реального времени. Поэтому цель настоящей работы – реализовать модель прогнозирования на основе штрафных сплайнов, предоставляющую возможность прогнозирования в масштабе реального времени и обеспечить высокую оценку качества прогнозирования, определяя оптимальные входные параметры сплайна без высоких потерь в скорости получения прогноза.

1 АНАЛИЗ ПРЕДМЕТНОЙ ОБЛАСТИ

1.1 Прикладные области краткосрочного прогнозирования

Прогнозирование – это одна из важнейших задач, которая охватывает множество областей, включающих бизнес и промышленную отрасль, государственный сектор, экономику, науки об окружающей среде, медицину, социологию, политику и финансы. В большинстве случаев задача прогнозирования связана с данными, представленными в виде временного ряда [1]. В статистике, экономике, обработке сигналов или результатов физических экспериментов, а также во многих других областях под временным рядом понимаются последовательно измеренные через некоторые временные промежутки данные. Прогнозирование временных рядов заключается в построении модели для предсказания будущих событий на основе уже полученных предыдущих значений. В зависимости от количества прогнозируемых значений выделяют краткосрочное, среднесрочное и долгосрочное прогнозирование. Задача краткосрочного прогнозирования состоит в определении значений на 1-3 шага вперед [1] и используется во множестве прикладных областей.

Прежде всего, краткосрочное прогнозирование актуально для предсказаний различных природных явлений, таких как температура воздуха, уровень воды, количество осадков, значение солнечной радиации и т.д. В дальнейшем эти значения могут использоваться для корректировки прогноза погоды, контроля стихийных бедствий и для развития возобновляемых источников энергии. Так, например, в [2] утверждается, что «стохастический и прерывистый характер ветровой энергии создает повышенный риск надежности системы и качества электроэнергии» и что одним из решений этой проблемы является использование краткосрочного прогнозирования ветровой генерации с целью уменьшения ее неопределенности.

Помимо этого, краткосрочное прогнозирование используется и в традиционной энергетике, в основном, для контроля загруженности

энергосистемы, с целью экономии ресурсов и предупреждения аварий. Отдельно в системе электроснабжения разрабатываются методы прогнозирования для определения ожидаемых электрических нагрузок для увеличения выгоды эксплуатации системы.

С экономической точки зрения краткосрочное прогнозирование применяется для предсказания значений спроса, значений цен, а также в задаче прогнозирования финансовых рынков. Инвесторы в финансовые активы заинтересованы в прогнозировании их доходности, в других случаях, инвестиционные решения могут приниматься относительно прогнозирования процентов тарифов, опционов и курсов валют [1]. На государственном уровне краткосрочное прогнозирование эффективно в задаче определения будущих значений макроэкономических показателей. В [3] отмечается, что, например, прогнозирование инфляции имеет решающее значение не только для коммерческих банков, но и для государственного аппарата.

Методы прогнозирования используются в медицине и фармацевтике. Так, краткосрочный прогноз на несколько недель вперед применяется в оперативном управлении и при выявлении эпидемических вспышек заболеваемости [4]. В клинической медицине прогноз трактуется как предвидение вероятного развития и исхода заболевания, то есть оценивает возможность возникновения болезни и возможность ее развития.

При наличии оперативной информации возможно использование краткосрочного прогноза для предсказания чрезвычайных ситуаций техногенного и природного характера. Как отмечено в [5] при использовании своевременно поступающей информации о возможных предвестниках чрезвычайной ситуации или о возникающих рисках возможен краткосрочный прогноз места, силы и времени наступления чрезвычайной ситуации.

В сельском хозяйстве и природопользовании также используются методы краткосрочного прогнозирования. В основном, это связано или с определением социально-экономического положения прогнозируемого

производства, или для контроля состояния природных запасов, сельскохозяйственной продукции.

1.2 Обзор методов краткосрочного прогнозирования

Существует множество методов краткосрочного прогнозирования. Исходя из [6], все методы разделяются на две группы в зависимости от степени их формализации: интуитивные и формализованные. Интуитивные методы применяются, если объект слишком прост или, наоборот, сложен настолько, что не представляется возможности учесть аналитически все факторы, влияющие на прогноз. К ним относятся, как индивидуальные экспертные оценки, включающие в себя метод «интервью», сценариев и т.д., так и коллективные, к которым относятся метод анкетирования, эвристического прогнозирования, «комиссий» и некоторые другие.

Формализованные методы, в свою очередь, включают в себя статистические и структурные модели прогнозирования [7] (или обозначаемые в [8], как методы, основанные на искусственном интеллекте).

Статистические или математические модели описывают зависимость между временным рядом и внешними факторами аналитически и включают в себя следующие группы: регрессионные модели, авторегрессионные и модели экспоненциального сглаживания [7] (в [4] последние к статистическим методам прогнозирования не относятся, а включены в отдельную группу методов прогнозирования на базе фильтрации).

Структурные модели представляют функциональную зависимость в виде некоторой структуры (например, графа или таблицы). К ним можно отнести нейросетевые модели, Байесовские сети, нечеткую логику и т.д.

Классификация методов прогнозирования по [7] представлена на рисунке 1:

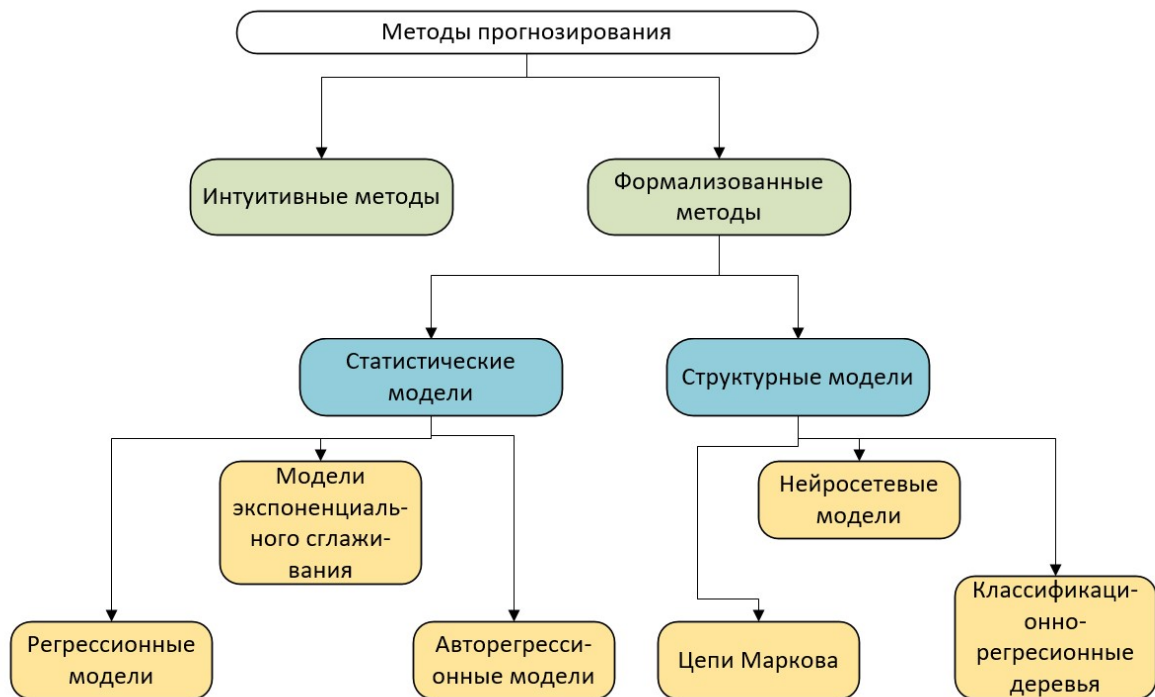


Рисунок 1 – Классификация методов прогнозирования по [7]

1.2.1 Основные статистические методы прогнозирования

Модели экспоненциального сглаживания относятся к адаптивным моделям прогнозирования, то есть могут изменять свою структуру и параметры, приспособившись к изменению условий. Экспоненциальное сглаживание ряда осуществляется по рекуррентной формуле

$$S_t = \alpha x_t + \beta S_{t-1}$$

где S_t – значение экспоненциальной средней в момент t ; α – параметр сглаживания ($0 < \alpha < 1$), $\beta = 1 - \alpha$.

Параметр α определяет зависимость прогноза от предыдущих данных. Повышение α обеспечивает увеличение веса наблюдений, поступивших позднее, а уменьшение значения параметра позволяет сглаживать случайные отклонения. [9] На рисунке 2 показано, как меняются экспоненциальные средние в зависимости от значения параметра сглаживания. Также экспоненциальное сглаживание всегда требует предыдущего значения экспоненциальной средней. Поиск начального значения можно разрешить

несколькими способами, подробно рассмотренными в [10], например, при наличии прошлых данных приравнять расчетное значение фактическому или вычислить среднее арифметическое имеющихся точек, а при их отсутствии - получить стартовые значения из процедуры «обратного прогноза».

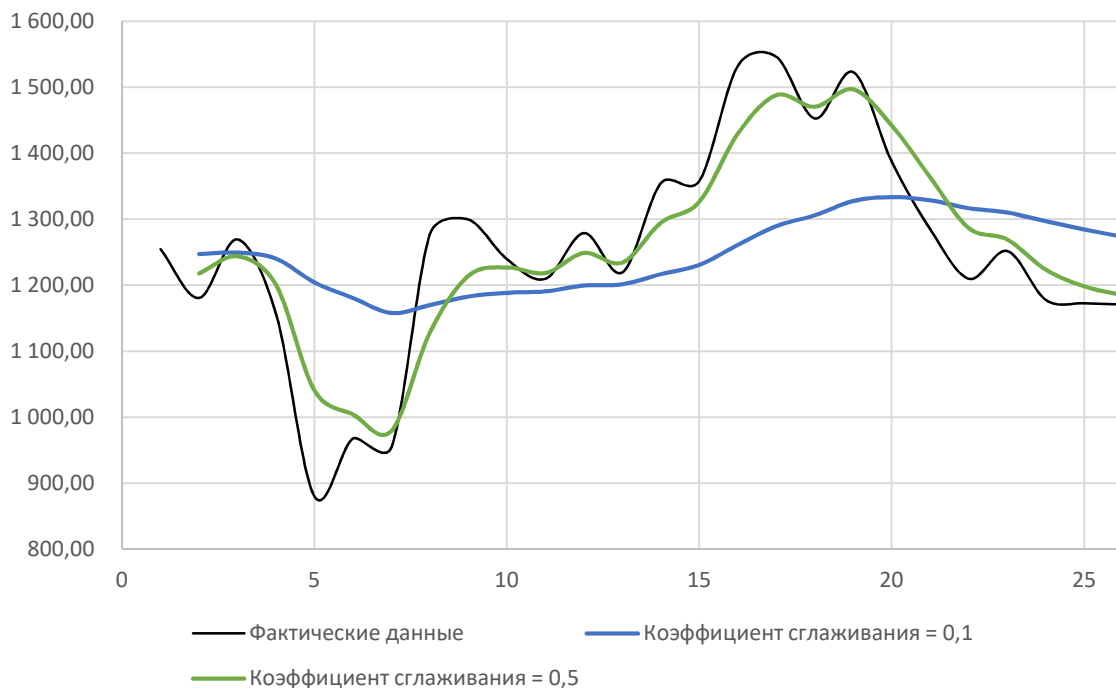


Рисунок 2 – Экспоненциальное сглаживание временного ряда

Экспоненциальная средняя дает систематическую ошибку, когда временной ряд имеет тенденцию линейного роста. Поэтому для анализа процессов, имеющих тренд, используется двойное экспоненциальное сглаживание или модель Хольта, в которой каждая составляющая процесса (уровень и тренд) сглаживаются отдельно

$$\begin{aligned}\Omega_t &= \alpha Y_t + (1 - \alpha)(\Omega_{t-1} - T_{t-1}), \\ T_t &= \beta(\Omega_t - \Omega_{t-1}) + (1 - \beta)T_{t-1}, \\ \hat{Y}_{t+p} &= \Omega_t + pT_t\end{aligned}$$

где α – коэффициент сглаживания уровня, а β – коэффициент сглаживания тренда. Третье уравнение определяет прогноз на p отсчетов времени вперед.

Главный недостаток подобных моделей в том, что временной ряд рассматривается изолированно от других явлений, которые так же могут

оказывать влияния на прогнозируемые значения. Например, многие явления характеризуются периодически повторяющимися сезонными эффектами, отсюда и временной ряд, отражающий подобное явление, будет иметь периодические сезонные колебания. [9]

Модель Винтерса или тройное экспоненциальное сглаживание расширяет модель Хольта тем, что включает в себя, помимо ряда общего уровня и тренда, сезонную составляющую

$$\begin{aligned}\Omega_t &= \alpha \frac{Y_t}{S_{t-s}} + (1-\alpha)(\Omega_{t-1} - T_{t-1}), \\ T_t &= \beta(\Omega_t - \Omega_{t-1}) + (1-\beta)T_{t-1}, \\ S_t &= \gamma \frac{Y_t}{\Omega_t} + (1-\gamma)S_{t-s}, \\ \hat{Y}_{t+p} &= (\Omega_t + pT_t)S_{t-s+p}\end{aligned}$$

где первое уравнение представляет сглаженный уровень без учета сезонной составляющей, второе – сглаженный тренд, а третье – сезонную составляющую. [7]

Регрессионные алгоритмы прогнозирования основаны на регрессионном анализе – методе, исследующим влияние независимых переменных X_1, X_2, \dots, X_n – регрессоров, на зависимую Y . Зависимая переменная представляет собой искомое посредством прогнозирования значение, а переменные-регрессоры – значения внешних факторов, которые оказывают существенное влияние на исход прогноза. Как правило, используются простые зависимости, которые легко интерпретируются и применимы на практике.

Самым простым вариантом регрессионной модели является линейная регрессия. Множественную регрессию в линейном случае можно представить в виде

$$y_j = \sum_{i=1}^n \alpha_i x_{ij} + \varepsilon_j$$

где α_i – коэффициенты модели; y_j, x_{ij} – значения j -й зависимой переменной и относящейся к ней i -й независимой переменной ($i = \overline{0, n}$); ε_j – случайная ошибка; n – число независимых переменных в модели. Определенную сложность в данной модели представляет необходимость знать значения всех факторов x_{ij} , что на практике почти невыполнимо. [7]

При оценке параметров уравнения регрессии обычно используется метод наименьших квадратов. Сущность МНК состоит в минимизации суммы квадратических отклонений между наблюдаемыми и расчетными величинами.

В классическом варианте МНК все исходные данные представляются равноценными между собой. Но наибольшую ценность для составления прогноза несут данные, полученные в результате более поздних наблюдений, поэтому целесообразно произвести дисконтирование – уменьшение ценности более ранней информации. [6]

Метод наименьших квадратов прост в реализации и подходит для использования в краткосрочном прогнозировании. Но наличие в данных наиболее сильно отклоняющихся значений точек оказывает существенное влияние на значения параметров уравнения.

Авторегрессионные модели основаны на том, что текущее значение прогнозируемой переменной задается функцией от прошлых значений этой же переменной. Модель авторегрессии порядка p можно представить в виде

$$y_t = \sum_{j=1}^p a_j \varepsilon_{t-j} + \delta_t$$

где α_j - коэффициенты авторегрессии, а δ_t - белый шум.

Наибольшее развитие модель авторегрессии получила в методологии, предложенной Боксом и Дженкинсом в середине 90-х годов, описанной ими в [11]. Из всего семейства алгоритмов и строящихся на их основе моделей самым известным и применимым на практике является алгоритм ARIMA (autoregressive integrated moving average). В качестве входных данных в нем

используются не значения временного ряда, а их разность порядка d , поэтому алгоритм подходит для нестационарных временных рядов, при условии, что число взятых разностей позволит достичь стационарности. Помимо параметров авторегрессии данный алгоритм включает в себя параметры скользящего среднего. Модель со скользящим средним порядка q имеет вид

$$y_t = \sum_{j=0}^q b_j \varepsilon_{t-j},$$

где b_j – параметры модели, а ε_t – ошибка модели.

Как отмечено в [12] в общем случае методика построения моделей типа ARIMA включает четыре этапа:

1) Идентификация стохастической модели – делается предварительный вывод о порядке интеграции d , а также о значениях порядков операторов авторегрессии и скользящего шума (обозначаются как p и q , соответственно).

2) Оценивание стохастической модели – определяются начальные оценки авторегрессии и начальные оценки параметров скользящего среднего, вычисляются оценки наименьших квадратов для соответствующих параметров стохастической модели.

3) Диагностика стохастической модели – диагностике подлежит качество модели и некоррелированность остатков.

4) Прогнозирование с помощью стохастической модели.

На сегодняшний день алгоритм ARIMA и прочие авторегрессионные модели и методы используются во множестве предметных областей при решении задач прогнозирования, так как являются довольно простыми в применении и прозрачны на всех этапах моделирования. Однако к недостаткам моделей данного класса можно отнести наличие большого количества параметров модели, которые сложно идентифицировать однозначно, а также отсутствие способности анализа нелинейных процессов, которые довольно часто встречаются на практике.

Модель ARIMA имеет множество модификаций, например, модель ARIMAX включает в себя некоторые экзогенные факторы X, наличие сезонности учитывается в моделях SARIMA – сезонного авторегрессионного скользящего среднего. Если же в прогнозировании используется несколько временных рядов, то они трактуются как векторы и исследуются с использованием модели VARIMA.

1.2.2 Использование нейронных сетей в задаче прогнозирования

Одним из методов, использующих машинное обучение и применяющихся в прогнозировании, является использование искусственных нейронных сетей. Их использование в задачах прогнозирования обладает рядом преимуществ. Прежде всего, в отличие от ранее рассмотренных методов нейронные сети не имеют ограничений на характер входных данных. Помимо временного ряда, в качестве входной информации могут служить данные о различных явлениях, влияющих на результаты прогнозирования. Еще одной отличительной чертой нейронной сети является ее нелинейность, то есть в процессе прогноза она способна устанавливать нелинейные зависимости между будущими и фактическими значениями процессов.

Существует множество вариантов применения искусственных нейронных сетей в краткосрочном прогнозировании. Нейронные сети могут отличаться архитектурой, количеством слоев и нейронов, использованием определенной функции активации или способом обучения.

К основным, используемым в прогнозировании видам нейронных сетей относятся:

1) Сети прямого распространения, к которым относятся многослойные и однослойные персептроны. В сетях прямого распространения отсутствуют обратные связи, т.е. входной сигнал последовательно преобразуется в выходной, проходя через некоторое количество слоев, при этом число нейронов на каждом слое может быть любым. Использование многослойного персептрона очень распространено в задачах прогнозирования. В частности,

для моделирования и прогнозирования часто используются сети прямого распространения с двумя слоями (один скрытый слой и один выходной). [13]

2) Рекуррентные нейронные сети – сети, в которых нейроны могут получать информацию не только от предыдущего слоя, но и от самих себя. Эта важная особенность позволяет моделировать нелинейные динамические системы, однако, простые рекуррентные сети (RNN) имеют существенный недостаток, заключающийся в быстрой потере информации с течением времени. Эта проблема решается в нейронных сетях с долгосрочной памятью (Long/Short Term Memory – LSTM), где используются фильтры (входной, выходной и забывающий) для защиты информации и клетка памяти, куда поступает информация из входного фильтра, а также в сетях с управляемыми рекуррентными нейронами (Gated Recurrent Unit – GRU), которые являются вариацией LSTM с использованием всего двух фильтров – фильтра обновления и фильтра сброса и с отличной реализацией связей (схемы описанных рекуррентных сетей представлены на рисунке 3). Как отмечено в [14], при прогнозировании количества солнечной энергии на сутки вперед эмпирические результаты показывают, что LSTM превосходит большое количество альтернативных методов.

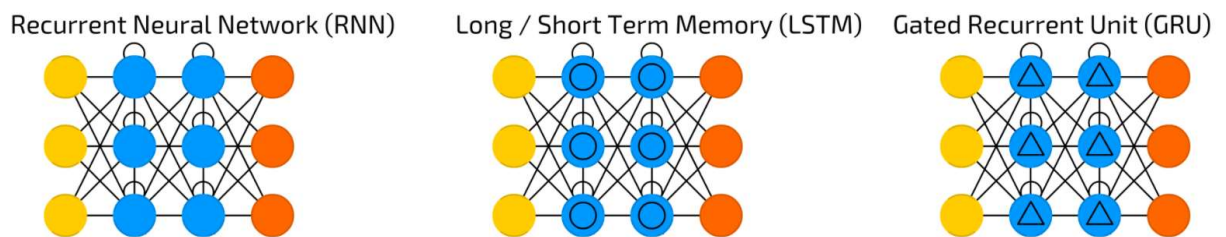


Рисунок 3 – Схемы рекуррентных нейронных сетей

3) Сверточные нейронные сети и глубокие сверточные нейронные сети используются, в основном, в задаче обработки изображений. Это однонаправленные многослойные сети, чередующие в своей архитектуре сверточные слои и слои подвыборки. Суть операции свертки, заключается в том, что каждая часть исследуемых данных (обычно, изображения) умножается на матрицу свертки поэлементно, затем полученный результат

суммируется и записывается в аналогичную позицию получаемого на выходе изображения. Несмотря на довольно узкую направленность, методы, основанные на сверточных нейронных сетях, используются и в прогнозировании [15,16]. Более того, в [15] отмечается, что точность и стабильность прогноза, основанного на модели сверточных нейронных сетей, в условиях прогнозирования в реальном времени, выше, чем у прогнозов, основанных на нейронных сетях обратного распространения и анализе временных рядов.

Помимо представленных выше, существует множество других видов архитектур нейронных сетей, таких как сети Конохена, автокодировщики, нейронные сети Хопфилда и т.д., но в их основные области применения на сегодняшний день не входит краткосрочное прогнозирование, поэтому подробно рассматриваться они не будут.

1.2.3 Использование сплайн-функций в задаче прогнозирования

Один из отличительных подходов к задаче прогнозирования описан в [16] и представляет собой использование аппроксимирующих кусочно-полиномиальных функций, заменяющих регрессионные соотношения при моделировании и анализе процессов. При этом выбор останавливается на таком виде кусочно-полиномиальных функций как сплайн-функции. В [18] отмечено, что особенностью сплайнов является то, что они состоят из отрезков степенного полинома малого порядка (степени), которые сопряжены (склеены) в заданных узловых точках процесса (узлах решетчатой функции).

Существуют различные виды сплайнов, применяемых в разнообразных прикладных задачах: регрессионные сплайны, В-сплайны, штрафные сплайны, сглаживающие сплайны и т.д.

Сглаживающие сплайны представляют собой метод аппроксимации исходных зашумленных данных. Наиболее известным примером сглаживающего сплайна является кубический сглаживающий сплайн. Его характеристика подробно описана в [18].

Регрессионные сплайны отличаются от сглаживающих тем, что данные аппроксимируются определенным набором базисных сплайн-функций с уменьшенным количеством узлов, обычно с использованием метода наименьших квадратов. Штрафы за негладкость отсутствуют.

Штрафные сплайны (P-splines) представляют собой сочетание уменьшенного количества узлов аналогично регрессионным сплайнам и штрафами за негладкость сглаживающих сплайнов, которые подавляют нелинейность. В [19] сравнивались модели прогнозирования, основанные на штрафных сплайнах и кубических сглаживающих сплайнах. Результаты показали, что согласно значениям среднеквадратической ошибки (Mean Squared Error – MSE), среднеабсолютной ошибки (Mean Absolute Error – MAE) и среднеабсолютной процентной ошибки (Mean Absolute Percentage Error – MAPE) для временного ряда, лучший результат показывает модель, основанная на штрафных сплайнах.

2 РАЗРАБОТКА МЕТОДА ПРОГНОЗИРОВАНИЯ НА ОСНОВЕ ШТРАФНОГО P-СПЛАЙНА

Одной из основных задач современного краткосрочного прогнозирования является задача уменьшения времени, затрачиваемого на прогнозирование без ухудшения качества прогноза. Особенно это необходимо при разработке прогнозных моделей, работающих в масштабе реального времени. Предложенная в данной работе математическая модель штрафного сплайна предполагает его построение на основе поступающих в реальном времени данных и последующее вычисление прогнозируемых значений. То есть описанный подход выполняет задачу формирования прогноза в режиме реального времени одновременно с обеспечением качества точности прогнозирования.

2.1 Расчетная схема и режимы функционирования штрафного P-сплайна

Используется расчетная схема сплайна, представленная в [20]. Отметим некоторые ключевые особенности этой схемы.

Для реализации режима реального времени используется рекуррентная форма расчета коэффициентов сплайна для каждого i -го звена на основании h отсчетов:

$$S_i(\tau) = a_0^i + a_1^i\tau + a_2^i\tau^2 + a_3^i\tau^3, \quad -q \leq \tau \leq h - q, \quad (1)$$

где q – номер отсчета t_j^i внутри i -го звена ($j = \overline{0, h}; q = \overline{0, h-1}$), в котором сопряжены непрерывные производные сплайна; $\tau = t_j^i - t_0^i$ – номер отсчета t_j^i внутри i -го звена, в котором сплайн вычисляется как $\tau = \overline{-q, h - q}$: $h = (t_h^i - t_0^i) / \Delta t$ – количество отсчетов в i -м звене сплайна.

Экстремальный функционал для предлагаемого сглаживающего сплайна имеет вид:

$$J(S) = (1 - \rho)(h\Delta t)^2 \int_{t_0^i}^{t_h^i} [S''(t)]^2 dt + \rho \sum_{j=0}^h [S(t_j^i) - y(t_j^i)]^2 \quad (2)$$

где ρ – весовой множитель, определяющий параметры равновесия между сглаживающими и интерполяционными свойствами сплайна $\rho \in [0, 1]$; Δt – интервал дискретизации измерений; $y(t_i) = f(t_i) + \xi(t_i)$ – измерения, представленные суммой полезного сигнала и случайной погрешности, соответственно.

Разрывные коэффициенты находятся из условия минимизации экстремального функционала (2). Для дефекта $d = 2$ – это a_2^i и a_3^i .

В зависимости от расположения точки q для i -го звена и требуемых моментов вычисления сплайна τ возможно несколько режимов функционирования сплайна. Все режимы рассмотрены в [21], основываясь на этом, приведем основную информацию по каждому из них.

Последовательный режим функционирования изображен на рисунке 4. Предполагается, что измерения $y(t_j)$ поступают для обработки группами из h наблюдений. Обработка также ведется для всей группы $\tau \in [0, h]$. Сопряжение непрерывных производных сплайна осуществляется в момент времени $t_h^{i-1} = t_0^i$. Время запаздывания результата Θ для момента времени t_τ^i складывается в этом случае из двух слагаемых, а именно собственного времени вычисления сплайна Θ_0 и времени окончания сбора h наблюдений i -го звена ($\Theta = \Theta_0 + (h - \tau)\Delta t$, $\tau = \overline{0, h}$).

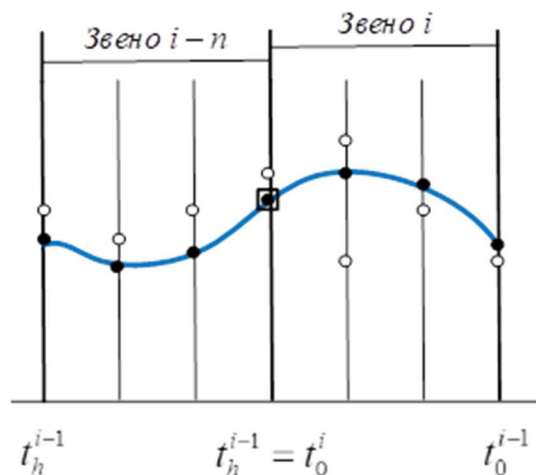


Рисунок 4 – Последовательный режим функционирования сплайна

Для прогнозирования в реальном масштабе времени используется текущий режим функционирования рекуррентного штрафного сплайна.

Вычисление производится при поступлении каждого измерения с использованием $(h - l)$ предыдущих измерений. В этом случае сопряжение осуществляется в момент времени $t_{q+1}^{i-1} = t_q^i$, $\tau = \overline{1, h-q}$. Запоздывание $\Theta = \Theta_0 + (h - q - \tau)\Delta t$, $\tau = \overline{0, h-q}$.

Как отмечено в [20] «формулы для оценки коэффициентов в таком режиме принимают наиболее простой вид».

Например, для $q = 0$ и дефекта $d = 2$:

$$\begin{aligned}
 a_0^i &= a_0^{i-1} + a_1^{i-1}(h - q) + a_2^{i-1}(h - q)^2 + a_3^{i-1}(h - q)^3; \\
 a_1^i &= a_1^{i-1} + 2a_2^{i-1}(h - q) + 3a_3^{i-1}(h - q)^2; \\
 a_2^i &= \frac{\rho(F_1^i C - F_2^i A)}{BC - A^2}; \\
 a_3^i &= \frac{\rho(F_2^i B - F_1^i A)}{BC - A^2};
 \end{aligned} \tag{4}$$

$$\begin{aligned}
 A &= 6(1 - \rho)h^3(h - 2q) + \rho H_5; \\
 B &= 4(1 - \rho)h^3 + \rho H_4; \\
 C &= 12(1 - \rho)h^3(h^2 - 3hq + 3q^2) + \rho H_6; \\
 F_1^i &= \sum_{k=0}^h y(t_k^i)(k - q)^2 - a_0^i H_2 - a_1^i H_3; \\
 F_2^i &= \sum_{k=0}^h y(t_k^i)(k - q)^3 - a_0^i H_3 - a_1^i H_4; \\
 H_n &= \sum_{k=0}^h (k - q)^n.
 \end{aligned} \tag{5}$$

Коэффициенты a_0^i, a_1^i найдены из условия сопряжения звеньев

$S^{(k)}(t_{q+1}^i)_+ = S^{(k)}(t_q^i)_-$, $k = 0, 1$, а коэффициенты a_2^i, a_3^i из условия минимизации

$$(2) \frac{\partial J(S)}{\partial a_2^i} = 0, \frac{\partial J(S)}{\partial a_3^i} = 0.$$

Текущий режим прогнозирования изображен на рисунке 5.

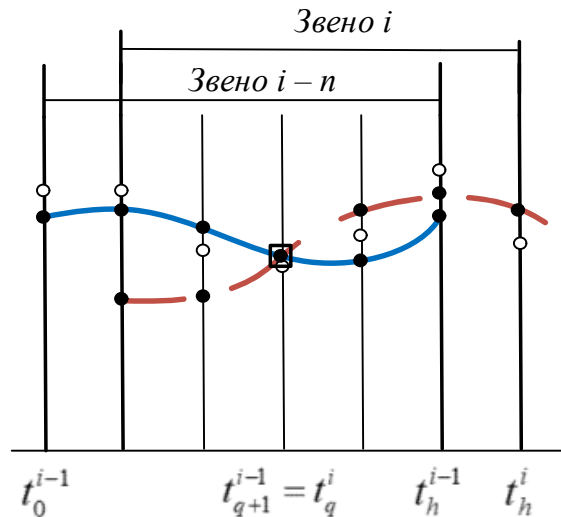


Рисунок 5 – Текущий режим прогнозирования

В режиме многократного сглаживания, как и в текущем режиме, вычисления производятся при поступлении каждого измерения, но с использованием $(h - 1)$ значений сплайна, вычисленных в моменты времени $t_{h-q}^{i-1}, t_{h-q}^{i-2}, \dots, t_{h-q}^{i-(h-1)}$. Запаздывание такое же, как и в текущем режиме.

Схематически режим многократного сглаживания изображен на рисунке 6.

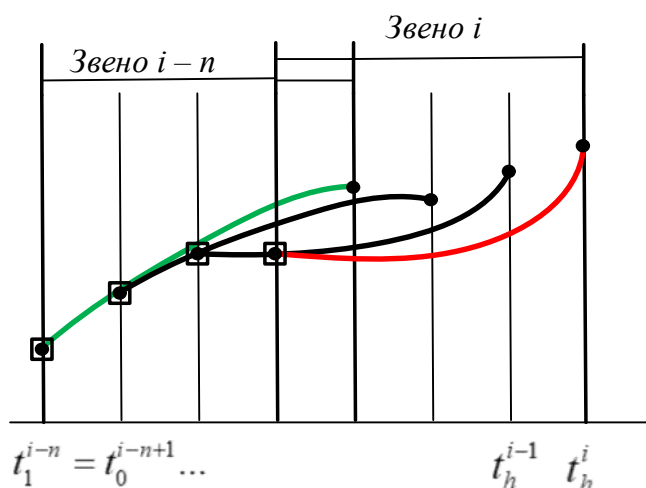


Рисунок 6 – Режим многократного сглаживания

2.2 Показатели качества прогнозирования

Для получения данных о точности прогноза и дальнейшего анализа результатов прогнозирования применяются такие показатели качества, как оценки погрешностей прогнозирования. При этом, существует несколько вариантов оценивания погрешностей, каждый из которых подходит под конкретную задачу или наиболее пригоден при использовании определенного типа входных данных.

Ниже рассмотрены некоторые виды погрешностей, предлагаемые для оценок качества прогнозирования в [19, 22].

Абсолютные погрешности прогнозирования основаны на вычислении остатка e_t :

$$e_t = (y_t - f_t^{(m)}),$$

где y_t – измеренное значение в момент времени t , $f_t^{(m)}$ – прогнозное значение в момент времени t , полученное в результате используемой модели прогнозирования m . В качестве погрешностей моделей прогнозирования особенно часто используются средняя абсолютная ошибка (Mean Absolute Error – MAE) и средняя квадратичная ошибка или среднеквадратичное отклонение (Root Mean Square Error – RMSE).

$$MAE = \frac{1}{n} \sum_{i=1}^n |e_i|$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (e^2)}$$

Однако, использование абсолютных погрешностей в оценке качества прогнозирования имеет ряд существенных недостатков. Прежде всего, неприменимость оценивания моделей на объектах с различными шкалами измерений из-за зависимости абсолютных погрешностей от этой шкалы. Кроме того, данный вид погрешностей, как отмечено в [23] реагирует на выбросы. Получить более устойчивую ошибку в данном случае можно при использовании медианы абсолютной ошибки (Median Absolute Error):

$$MdAE = \text{median}_{i=1, \dots, n} |e_i| ,$$

где *median* - операция нахождения медианы.

Процентные (в русскоязычной литературе они называются относительными) погрешности рассчитываются на основе значения p_t :

$$p_t = \frac{|e_t|}{y_t} .$$

К процентным ошибкам относятся: средняя относительная ошибка прогноза (Mean Absolute Percentage Error – MAPE), средняя квадратичная процентная ошибка (Root Mean Square Percentage Error – RMSPE), а также их медианы.

$$MAPE = \frac{1}{n} \sum_{i=1}^n 100 |p_i| = \text{mean}_{i=1, \dots, n} (100 |p_i|)$$

$$RMSPE = \sqrt{\text{mean}_{i=1, \dots, n} (100 |p_i|)^2}$$

$$MdAPE = \text{median}_{i=1, \dots, n} (100 |p_i|)$$

$$RMdSPE = \sqrt{\text{median}_{i=1, \dots, n} (100 |p_i|)^2}$$

Основным недостатком этой группы погрешностей является ошибка деления на ноль, возникающая в случае, если измеряемое значение равно

нулю. Помимо этого, значение погрешностей данной группы различно в зависимости от того, ниже или выше находится прогнозное значение относительно реального, что приводит к несимметричности оценок. Если во входных данных присутствуют значительные выбросы, то они достаточно сильно влияют на значения оценок погрешностей.

Симметричные погрешности рассчитываются на основании значения, определяемого по формуле:

$$s_t = \frac{|e_t|}{(y_t + f_t)}$$

В группу симметричных погрешностей включены: симметричная абсолютная процентная ошибка (Symmetric Mean Absolute Percentage Error – sMAPE) и медиана средней абсолютной процентной ошибки:

$$sMAPE = \frac{1}{n} \sum_{i=1}^n 200 |s_i| = \text{mean}_{i=1, n} (200 |s_i|)$$

$$sMdAPE = \text{median}_{i=1, n} (200 |s_i|)$$

Также существует модификация sMAPE – Modified sMAPE (msMAPE), в знаменателе которой находится дополнительный член для избежание возможных проблем, возникающих при делении на ноль.

$$msMAPE = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - f_i|}{(|y_i| + |f_i|) / 2 + S_i}$$

$$S_i = \frac{1}{i-1} \sum_{k=1}^{i-1} |y_k - \bar{y}_{i-1}|, \bar{y}_{i-1} = \frac{1}{i-1} \sum_{k=1}^{i-1} y_k$$

Однако, аналогично процентным погрешностям ошибки данной группы также подвержены влиянию выбросов.

Масштабированные погрешности симметричны и, в отличие от предыдущих ошибок, устойчивы к выбросам. Для их расчета используется величина q_i

$$q_i = \frac{|e_i|}{\frac{1}{n-1} \sum_{i=2}^n |y_i - y_{i-1}|}$$

К группе масштабируемых ошибок относят среднюю абсолютную масштабированную ошибку (Mean Absolute Scaled Error – MASE), медиану абсолютной масштабированной ошибки и среднеквадратическую масштабированную ошибку (Root Mean Square Scaled Error – RMSSE).

$$MASE = \text{mean}_{i=1, n} |q_i|$$

$$MdASE = \text{median}_{i=1, n} |q_i|$$

$$RMSSE = \sqrt{\text{mean}_{i=1, n} (q_i^2)}$$

Следует учесть, что если в интервале прогнозирования реальные значения, идущие подряд, равны, то возникает ошибка деления на ноль.

К относительным погрешностям относятся оценки, которые рассчитываются на основе отношения ошибок для анализируемой модели прогнозирования и базовой модели.

Так относительная средняя абсолютная ошибка (Relative Mean Absolute Error – RelMAE) вычисляется как отношение средней абсолютной ошибки для анализируемой модели прогнозирования и базовой модели. Аналогично, относительное среднеквадратичное отклонение (Relative Root Mean Square Error – RelRMSE), только оценки в отношении вычисляются по формуле среднеквадратичного отклонения.

$$RMAE = \frac{MAE}{MAE^*}$$

$$RRMSE = \frac{RMSE}{RMSE^*}$$

Критерии определяют число случаев, когда исследуемая модель прогнозирования превосходит базовую, но не определяют насколько.

Если оценивается качество прогноза временных рядов с одинаковыми единицами измерения, при этом данные были предварительно обработаны, то

целесообразно использовать абсолютные погрешности. Процентные ошибки не рекомендуются для использования на необработанных данных в связи с их несимметричностью. Симметричные оценки подходят для значений, область изменения которых лежит в положительной полуплоскости, и при отсутствии выбросов. Если же данные содержат выбросы, то следует применять масштабируемые оценки, при этом горизонт прогнозирования должен быть очень большим, чтобы избежать ошибки деления на ноль.

В дальнейшем в работе при анализе результатов прогнозирования будут использоваться процентные погрешности, в которых, для того чтобы избежать несимметричности будет использоваться диапазон между максимальным и минимальным значением исходных данных:

$$MSPE = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (s_i - y_i)^2} \cdot \frac{100}{\max(y) - \min(y)}, [\%]. \quad (6)$$

3 ОПТИМИЗАЦИЯ ПАРАМЕТРОВ СПЛАЙНА

Применяемый в прогнозировании штрафной сплайн имеет несколько параметров настройки, которые существенно влияют на его свойства. Этими параметрами являются: сглаживающий множитель ρ , число измерений звена сплайна h и число прогнозируемых точек m . [17]

Так на рисунке 7 изображены фрагменты сплайнов в задаче восстановления полезного сигнала исходной функции (изображена на графике в виде пунктирной линии). Шум функции – 10 %, сглаживающий множитель $\rho = 0,5$, значения числа измерений звена сплайна h равны 5, 10 и 20, соответственно. Видно, что даже изменение одного параметра сплайна существенно влияет на конечный результат.

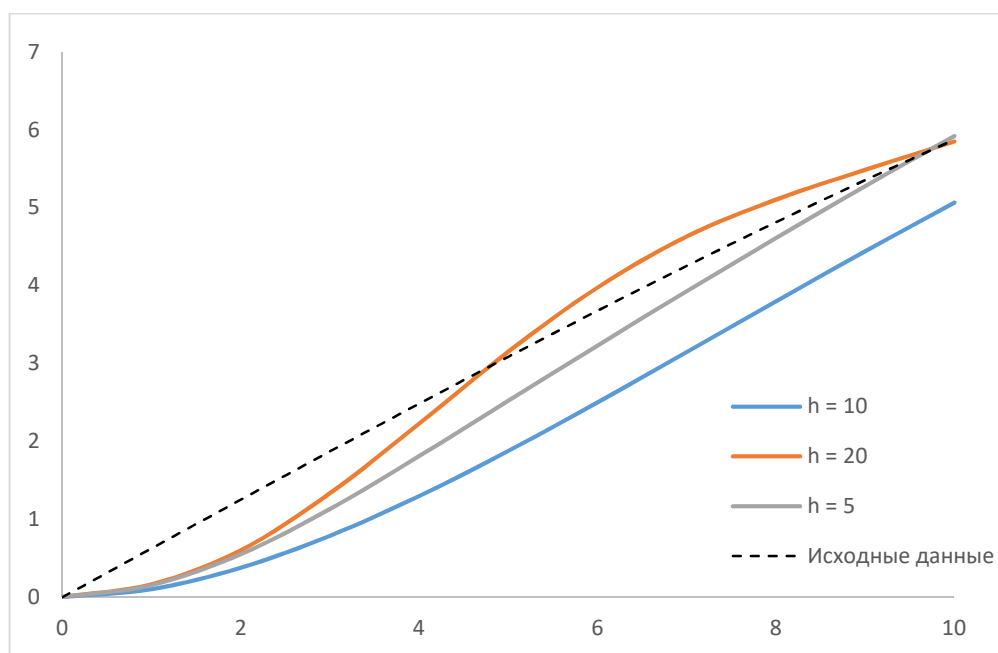


Рисунок 7 – Использование штрафного сплайна в задаче восстановления полезного сигнала

Для определения набора настраиваемых параметров, при использовании которых для конкретных входных данных результат прогнозирования будет иметь наименьшую погрешность, целесообразно использовать алгоритм оптимизации. В настоящее время существует множество оптимизационных методов, существенно различающихся по принципам работы. Поэтому для выбора подходящего для задачи определения настраиваемых параметров

сплайна оптимизационного алгоритма, необходимо рассмотреть возможные варианты оптимизации и провести сравнительный анализ для каждого на основе конкретных критериев.

Одним из самых важных критериев, необходимых к рассмотрению, является скорость выполнения. Так как прогнозирование на основе штрафных сплайнов предназначено так же и для формирования прогноза в масштабе реального времени, оптимизационный метод должен выполняться за максимально быстрое время, по сравнению с другими алгоритмами. Кроме этого, высокая скорость алгоритма не должна приводить к потерям в качестве прогноза. То есть выбранный алгоритм должен обеспечивать приемлемое соотношение скорости выполнения и качества прогнозирования. Помимо основных критериев так же необходимо проанализировать такие параметры оптимизационных методов, как сложность реализации метода и количество затрачиваемых ресурсов.

3.1 Сравнение методов оптимизации параметров сплайна

Как известно, задача оптимизации состоит в поиске экстремума целевой функции с заданными ограничениями. Ее математическая постановка дана в [24] и выглядит следующим образом: нужно определить значения вектора переменных $x = (x_1, x_2, \dots, x_m)$, которые удовлетворяют ограничениям вида:

$$g_i(x_1, x_2, \dots, x_m) \leq b_i$$

для всех $i = 1, \dots, k$ и при которых достигается максимум или минимум целевой функции $f(x_1, x_2, \dots, x_m)$:

$$f(x_1, x_2, \dots, x_m) \rightarrow (\max, \min).$$

В общем случае, большинство алгоритмов реализуют методы спуска к экстремуму функции таким образом, что значение целевой функции становится меньше вплоть до достижения экстремума. В зависимости от того, глобальный или локальный экстремум могут найти алгоритмы оптимизации, они подразделяются на алгоритмы глобального и локального поиска.

Алгоритмы локального поиска определяют один из возможных локальных экстремумов на множестве допустимых решений. Это может обеспечиваться как через случайный поиск, так и с использованием детерминированного спуска в область экстремума, среди которых различают методы нулевого порядка и градиентные, в которых используются частные производные соответствующего порядка.

К методам нулевого порядка (прямым методам) относятся: метод координатного спуска – поочередная оптимизация параметров вдоль осей одним из одномерных методов, координатного спирального спуска, метода Розенброка (метод вращающихся координат), метода Хука-Дживса, метода Нелдора-Мидта, метод оврагов и т.д.

Градиентные методы оптимизации представляют собой итерационные процессы оптимизации, направление поиска которых на каждом шаге совпадает с антиградиентом функции. [25] К ним относят: алгоритм наискорейшего спуска, метод Ньютона (алгоритм наискорейшего спуска с шагом, длина которого зависит от свойств минимизируемой функции), алгоритм стохастической оптимизации, а также нейронные сети, из которых наиболее используемыми в задаче оптимизации являются сети Хопфилда.

Алгоритмы поиска глобального экстремума направлены на поиск приближенного решения задачи оптимизации, то есть точки из множества оптимальных решений. В данном случае, поиск точного решения можно рассматривать как частный случай поиска приближенного решения. Они разделяются на детерминированные и статистические методы.

Детерминированные алгоритмы совершают поиск глобального решения посредством его поиска на всем допустимым множестве. Такими методами являются алгоритм Гомори или метод отсекающих плоскостей и интервальный метод ветвей и границ. Стохастические алгоритмы в отличие от детерминированных не теряют эффективность с возрастанием условия задачи, к ним относят: эволюционные или генетические алгоритмы, поведенческие стратегии, машину Больцмана и т.д.

Достоинства и недостатки наиболее популярных оптимизационных алгоритмов приведены в приложении А. Исходя из них, можно сделать вывод, что генетический алгоритм представляет собой метод оптимизации, в котором наиболее удачно соотносятся скорость выполнения и качество получаемых результатов, при этом на соотношение этих критериев можно влиять, изменяя размер итераций. Также использование генетических алгоритмов обусловлено тем, что они работают на различных входных данных без предварительной обработки.

3.2 Использование генетического алгоритма для задачи оптимизации

Генетический алгоритм использует для работы эволюционные принципы наследственности, изменчивости и естественного отбора. [26] Общая схема генетического алгоритма приведена на рисунке 8.

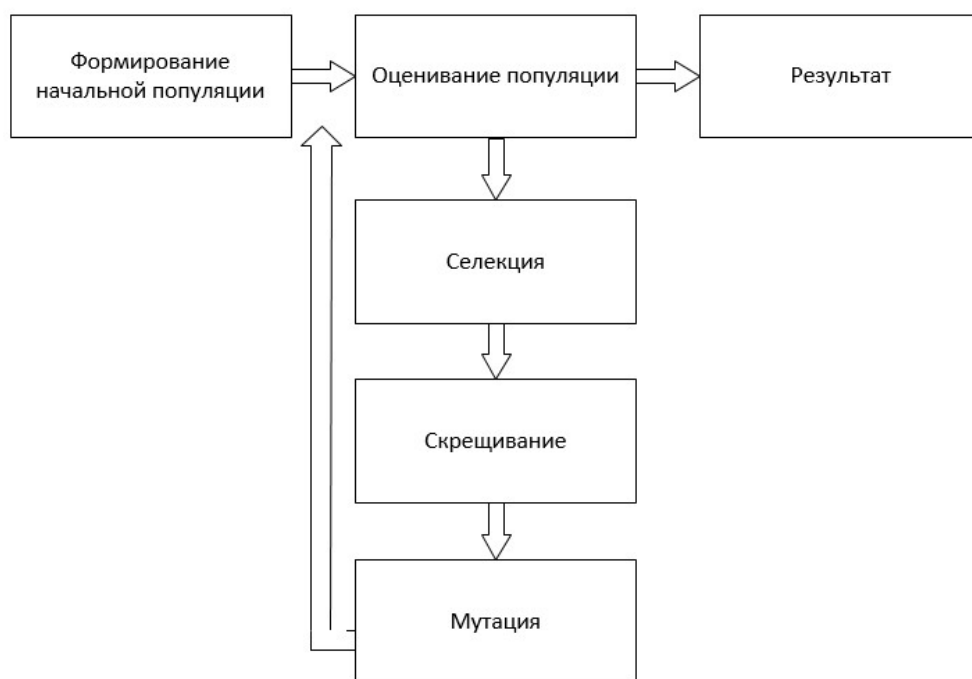


Рисунок 8 – Общая схема генетического алгоритма

Генетический алгоритм работает с популяцией особей, в хромосоме (генотип) каждой из которых закодировано возможное решение задачи (фенотип).

При запуске алгоритма, прежде всего, формируется случайная популяция, обычно случайным образом. Далее происходит этап оценивания популяции, где для оценки качества закодированных решений используется функция приспособленности или целевая функция. По результатам оценивания особей из поколения выбираются самые приспособленные для дальнейшего скрещивания, то есть происходит этап селекции особей. Существует множество вариантов селекции, ниже приведены наиболее используемые способы.

Вероятность i -ой особи принять участие в скрещивании в рулеточной селекции вычисляется по формуле:

$$p_i = \frac{f_i}{\sum_j f_j},$$

где f_j - значение приспособленности i -ой функции. Название данного варианта селекции связано с тем, что процесс отбора для скрещивания напоминает игру в «рулетку», где площадь каждого сектора рулеточного круга пропорциональна значению вероятности p_i .

При использовании для отбора особей селекции усечением после вычисления значений функции приспособленности для каждой особи выбираются ln лучших особей, где l – значение порога отсечения, $l \in [0,1]$, n – размер популяции. Обычно значение порога отсечения выбирается в интервале от 0,3 до 0,7, в зависимости от поставленной задачи.

Еще одним известным способом селекции особей является турнирный отбор. В случае его использования из популяции случайным образом выбираются t особей, из них самая приспособленная допускается к скрещиванию. Данная операция повторяется, пока не наберется необходимое для следующего поколения число особей.

Те особи, которые прошли этап селекции переходят на этап скрещивания. В результате скрещивания формируются хромосомы потомков,

состоящие из генетической информации, хранящейся в хромосомах родительских особей.

При целочисленном кодировании информации в хромосомы особей в качестве операторов скрещивания используют 1-точечный, 2-точечный и однородный операторы кроссинговера.

В случае 1-точечного и 2-точечного кроссинговера в двоичном представлении хромосом родителей случайно определяется одна или две точки разрыва, между которыми производится обмен частями родительских хромосом. При использовании однородного оператора кроссинговера разряды родительских хромосом чередуются независимо друг от друга. Для этого определяют вероятность p_0 , что i -ый разряд хромосомы первого родителя попадет к своему первому потомку, а разряд хромосомы второго родителя – к другому потомку. [26]

В результате скрещивания создаются потомки для формирования нового поколения. При этом совершенно не обязательно в новое поколение должны включаться только особи-потомки. Обычно, исходя из коэффициента T , который называют «разрыв поколений» или *generation gap*, часть поколения формируется из особей-потомков, а остальная часть заполняется элитными особями из предыдущего поколения. Данный способ может обеспечить скорость сходимости генетического алгоритма.

Этап мутации необходим для внесения случайных изменений в хромосомы особей. При этом, работа оператора мутации зависит от вероятности ее применения, которая может быть различной в зависимости от характера задачи. Использование мутации помогает при использовании алгоритма избегать попадания в локальные минимумы. В целочисленном кодировании наиболее популярным методом мутации является битовая мутация, где каждый бит хромосомы инвертируется с определенной вероятностью. Стоит отметить, что значение вероятности выбирается достаточно небольшим, чтобы не подвергать разрушению хромосомы с высокой функцией приспособленности.

Универсальность генетического алгоритма определяется тем, что только функция приспособленности и кодирование решений зависит от конкретной задачи. Все остальные шаги для любой задачи производятся одинаково.

В [26] перечислены необходимые этапы настройки генетического алгоритма. Исходя из них, при использовании генетического алгоритма в задаче оптимизации необходимо:

1. Определить количество оптимизируемых переменных и тип каждой, которые в дальнейшем будут закодированы в хромосоме.
2. Определить критерий оценки пользователей в виде функции приспособленности.
3. Выбрать способ кодирования и его параметры.
4. Определить параметры генетического алгоритма такие, как размер популяции, тип селекции, вероятности используемых генетических операторов, величина разрыва поколений.

Так как результатом выполнения генетического алгоритма должны быть значения регулируемых параметров штрафного сплайна, в каждой хромосоме необходимо хранить две переменные, одна из которых представляет значение h параметра и находится в диапазоне значений от 2 до 21, вторая – параметр ρ , диапазон которой от 0 до 1 (в реализации алгоритма используется значение от 0 до 100, чтобы работа велась с целочисленным кодированием).

В качестве критерия оценки особей принимается значение процентной погрешности (MSPE). Чем меньше значение погрешности, тем более приспособленной является особь.

Ниже приведены значения параметров генетического алгоритма, реализующего задачу оптимизации регулируемых параметров сплайна:

- размер популяции: 50 особей;
- тип селекции: селекция усечением, порог отсечения – 0,3;
- скрещивание: 1-точечный кроссинговер, вероятность скрещивания – 0,7;

- мутация: вероятность мутации – 0,042;
- разрыв поколений: 0,8.

Начальная популяция формируется из случайных значений в заданном диапазоне для каждой переменной. В дальнейшем при оценке популяции оценивается значение оценки качества прогнозирования с использованием данных параметров и массив особей сортируется в порядке от самой приспособленной до менее приспособленной. Функция приспособленности в данном случае имеет вид погрешности MSPE (6). Создается массив для нового поколения. Формируются случайным образом два родителя из тех особей, которые входят в значение выбранных селекцией с помощью порога отсечения. Если случайное значение от 0 до 1 меньше, чем значение вероятности скрещивания, то данные особи скрещиваются с использованием одноточечного кроссинговера. После этого сформированный потомок подвергается случайной мутации (произойдет или нет мутация зависит от значения вероятности мутации). Затем потомок попадает в массив нового поколения. Потомки заполняют поколение не полностью, так как значение Разрыва поколений меньше единицы. Оставшиеся места в новом поколении занимают элитные особи. Оценивание поколения и операции над ним совершаются до тех пор, пока не будет достигнуто максимальное число итераций. Для данного алгоритма наиболее оптимальным значением выбрано 20 итераций.

4 РАЗРАБОТКА ПРОГРАММНОГО ПРИЛОЖЕНИЯ

4.1 Среда разработки и язык программирования

Для разработки программного приложения выбран объектно-ориентированный язык программирования C#. Данный язык является членом семейства языков программирования, основанных на C (куда также входят C, Objective C, Java, C++) и разделяет с ними похожий синтаксис.

Кроме того, C# поддерживает множество средств, которые традиционно встречаются в различных языках функционального программирования (например, LISP или Haskell), например, лямбда-выражения и анонимные типы. Вдобавок, с появлением технологии LINQ (Language Integrated Query — язык интегрированных запросов), язык C# стал поддерживать конструкции, которые делают его довольно уникальным в мире программирования. Несмотря на все это, наибольшее влияние на него оказали именно языки, основанные на C. [27]

В качестве интегрированной среды разработки используется Microsoft Visual Studio 2017. Для разработки графического интерфейса используется Windows Presentation Foundation – система для построения клиентских приложений Windows с визуально привлекательными возможностями взаимодействия с пользователем. С помощью WPF можно создавать широкий спектр как автономных, так и размещенных в браузере приложений. [28]

Инфраструктура WPF специально создавалась для того, чтобы объединить различные несвязанные программистские задачи такие, как построение окон с элементами управления, поддержка двухмерной и трехмерной графики, поддержка потокового видео и документов нефиксированного формата, в единую объектную модель. Очевидное преимущество здесь в том, что WPF представляет собой единый симметричный API-интерфейс для всех потребностей, связанных с построением графических пользовательских интерфейсов. [27]

Для более четкого отделения внешнего вида приложения от программной логики, которая им управляет, в WPF-приложениях используется XAML, который определяет пользовательский интерфейс через разметку XML.

4.2 Структура программного приложения

При проектировании архитектуры приложений особенно важной частью является декомпозиция программы на подсистемы и организация их взаимодействия между собой и внешним миром. Эти подсистемы называются функциональными модулями. Модуль – это отдельная функционально осмысленная и законченная программная единица (подпрограмма), которая обеспечивает решение некоторой задачи. Стоит отметить, что таким образом в приложении реализуется многослойная архитектура. В ней каждый слой агрегирует ответственности и абстракции уровня, расположенного непосредственно под ним, или, если используется нестрогий многослойный стиль, всех более низких слоев. [29]

Для реализации программного приложения, обеспечивающего возможность краткосрочного прогнозирования временных рядов, были добавлены четыре модуля:

1) `DataModel` – включает в себя класс, описывающий предметную область. В данном случае, это данные, считываемые с файла пользователя и значения входных параметров сплайна.

2) `ForecastingModule` – библиотека классов, обеспечивающая парсинг пользовательских файлов и содержащая классы для прогнозирования данных на основе входной информации.

3) `OptimizationModule` – содержит классы, реализующие генетический алгоритм для использования его в задаче оптимизации входных параметров сплайна.

4) `DesktopUIModule` – пользовательский интерфейс.

Диаграмма пакетов (модулей), отображающая связи между ними, представлена на рисунке 9.

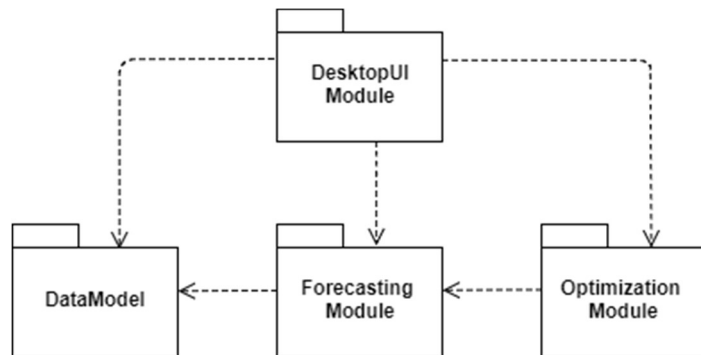


Рисунок 9 – Диаграмма модулей программного приложения

Взаимодействие классов между собой в программном приложении отображено на рисунке 10. Описание каждого класса представлено в приложении Б.

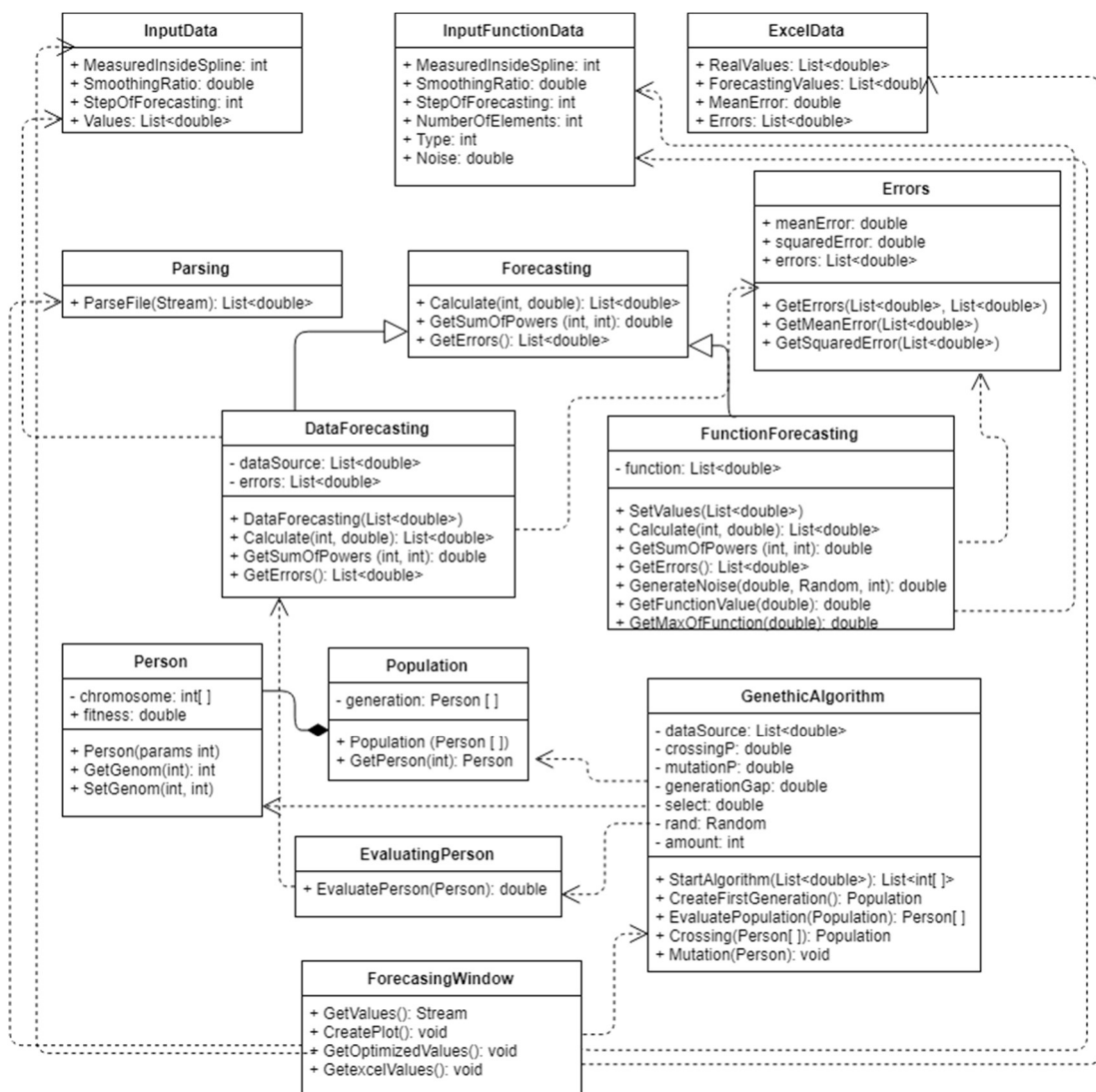


Рисунок 10 – Взаимодействие классов программного приложения

4.3 Функциональные требования к программному приложению

Функциональные требования определяют функциональность программной системы, которую необходимо построить, чтобы пользователи смогли выполнить свои задачи в рамках бизнес-требований [30].

Для отображения всех функциональных требований системы используется диаграмма вариантов использования на рисунке 11. Каждый вариант использования описывает с точки зрения действующего лица, группу действий в системе, которые приводят к конечному результату [31].

Пользователь может получать результаты прогнозирования, загружая в приложение собственные данные, предназначенные для прогноза, а также определяя значения входных параметров сплайна. Пользователь имеет возможность также экспортировать результаты в Excel. Кроме, этого пользователь может непосредственно исследовать прогнозные и сглаживающие свойства штрафного сплайна, без наличия собственных данных. Для этого ему необходимо определить параметры к одной из предоставленных тестовых функций и задать входные параметры сплайна.

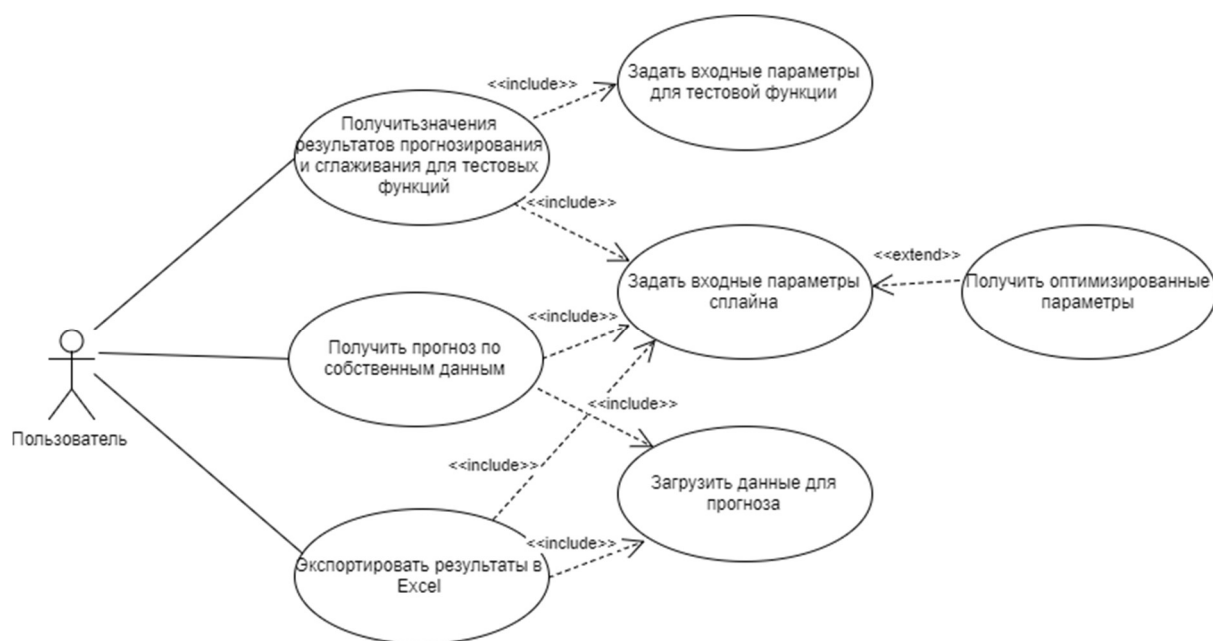


Рисунок 11 – Диаграмма вариантов использования для программного приложения

4.4 Интерфейс пользователя

В процессе определения требований к приложению было принято решение о создании однооконного пользовательского интерфейса. Такой вид интерфейса очень прост и удобен в управлении, в отличие от многооконных, он наглядно представляет все необходимые данные одновременно. Решение использовать десктоп-интерфейс обусловлено тем, что использование данного приложения не требует регистрации и авторизации пользователя и рассчитано на длительный период использования, например, в сфере образования, в научных исследованиях или в сфере финансов и экономики.

Прототип интерфейса представлен на рисунке 12.

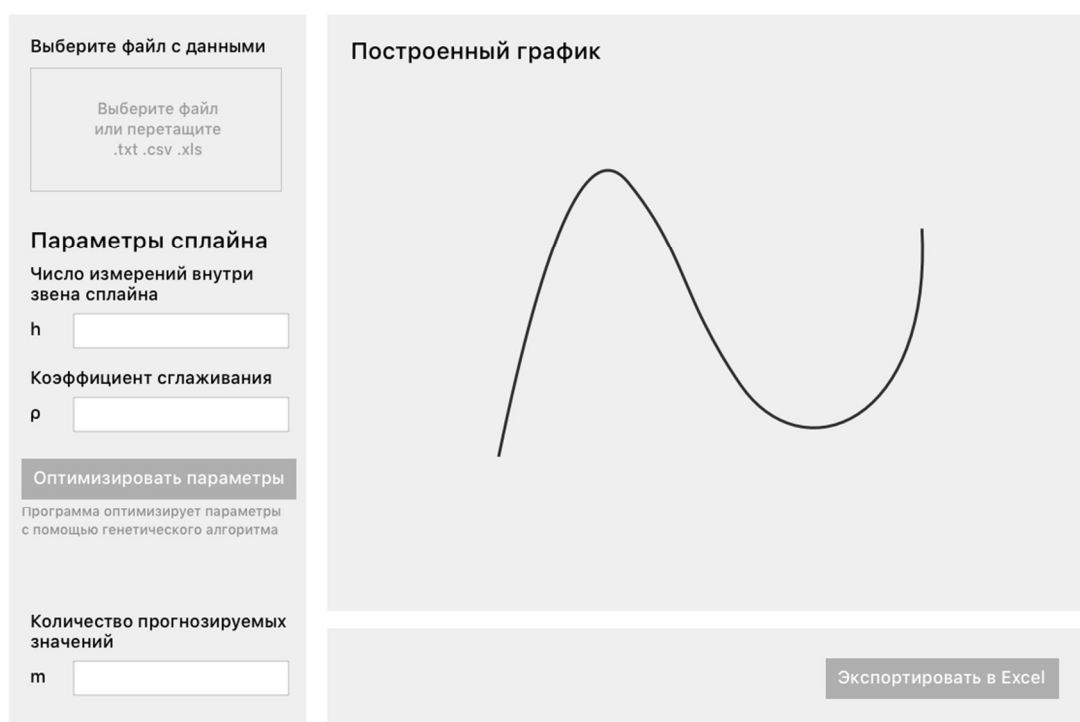


Рисунок 12 – Прототип интерфейса приложения

На рисунке 13 изображен экран программы. В левой половине окна расположен компонент для ввода входных данных. При нажатии на кнопку «Оптимизировать параметры» вместо ввода регулируемых параметров сплайна можно получить их оптимальные значения. В правой части экрана приложения расположен график функции, который строится на основе полученных пользовательских данных и входных параметров сплайна, а также

значение процентной погрешности и кнопка «Экспортировать в Excel», которая генерирует Excel-файл с данными погрешностей прогнозирования для каждой точки графика для более удобного дальнейшего анализа и использования этих данных.

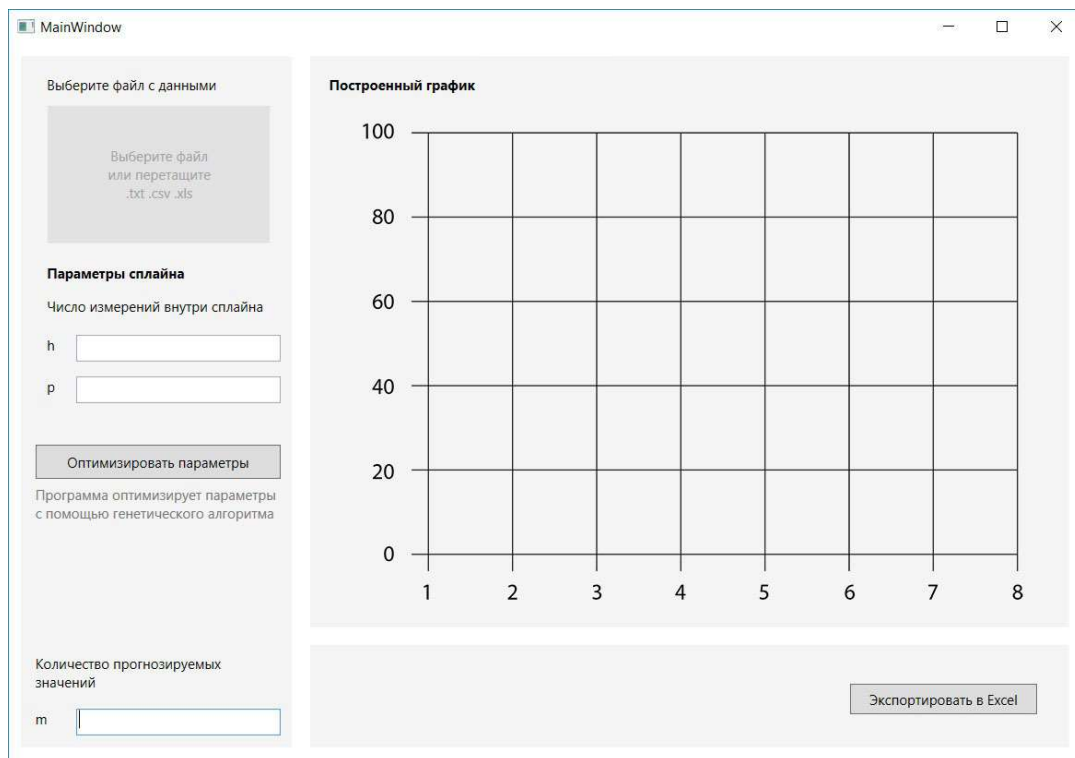


Рисунок 13 – Экран приложения

5 ИССЛЕДОВАНИЕ РАБОТЫ ПРОГРАММНОГО ПРИЛОЖЕНИЯ

5.1 Анализ результатов краткосрочного прогнозирования на основе штрафных P-сплайнов

Помимо использования сплайна конкретно для задач прогнозирования на пользовательских данных, прогнозные свойства сплайна анализировались на основе тестовых функций. Каждая из этих функций может подвергаться случайному зашумлению от 0 до 100 % от максимального значения функции. Их использование в задаче анализа качества прогнозирования помогает исследовать как прогнозные, так и сглаживающие свойства штрафного сплайна.

Так, на рисунке 14 представлен полезный сигнал функции $10\sin(\frac{2\pi x}{100})$, искаженный случайной помехой ее сигнал (шум равен 10 % от максимального значения функции) и полученный штрафной сплайн при входных параметрах $h = 10$, $\rho = 0,5$ и $m = 1$. Аналогичный график построен для функции $\sin(\frac{\pi x}{20})e^{(0,02x)} + 3$ (рисунок 15).

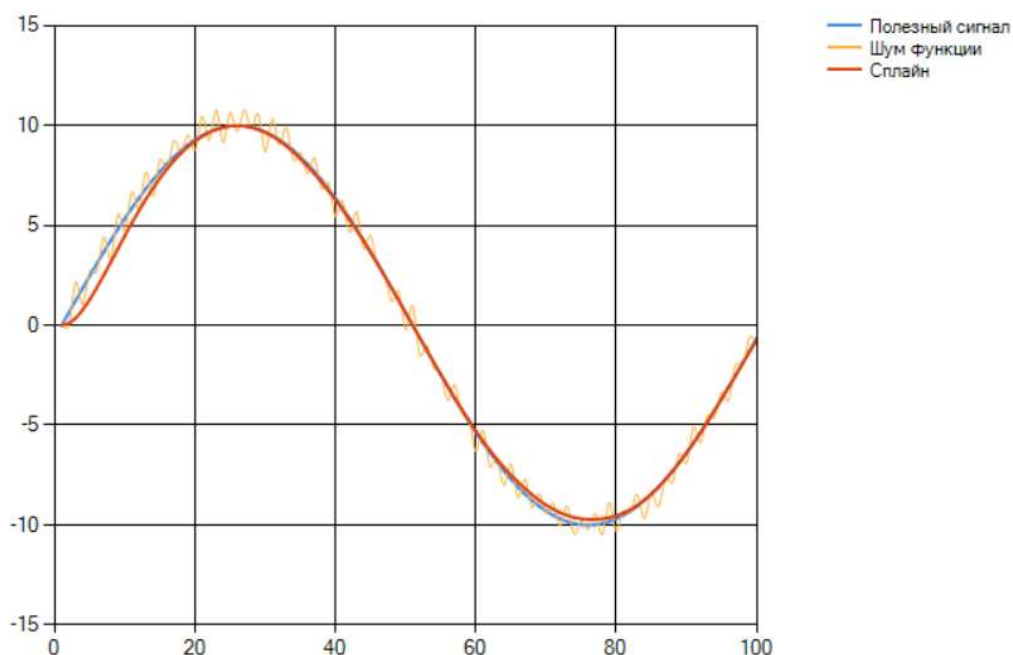


Рисунок 14 – Результат работы штрафного сплайна для тестовой функции

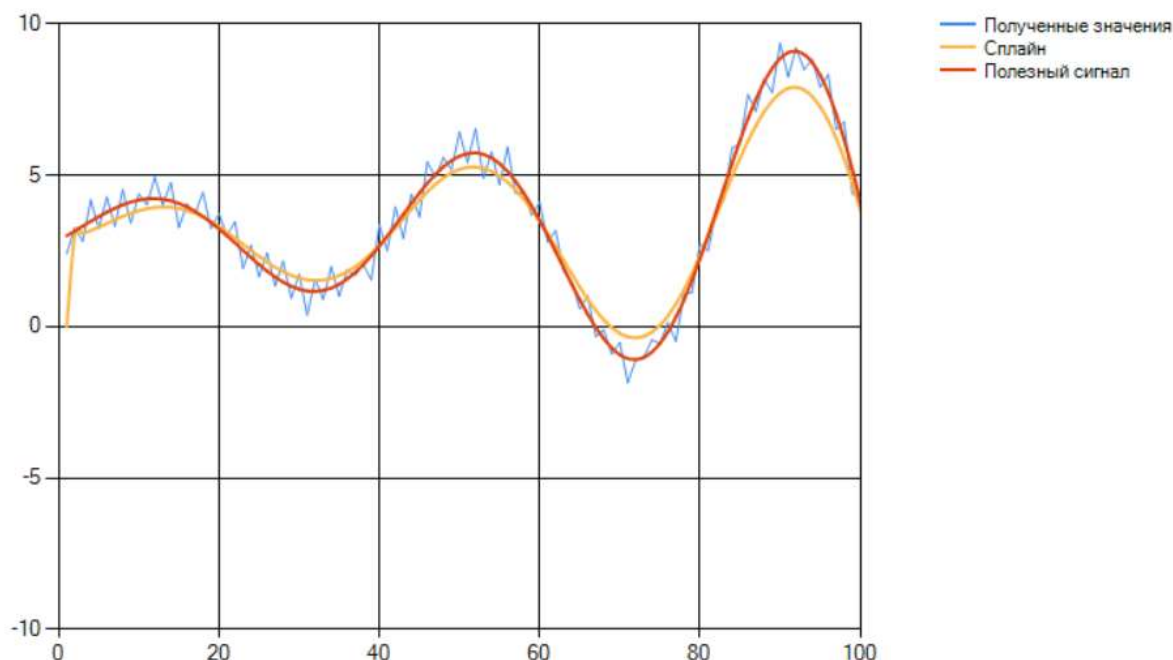


Рисунок 15 – Результат работы штрафного сплайна для тестовой функции

Для проверки восстанавливаемой функции вычисляется

$$MSPE = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (s_i - y_i)^2} \cdot \frac{100}{\max(y) - \min(y)}, [\%],$$

где s_i – значение полученной функции в данной точке, y_i – истинное значение функции в данной точке, n – число значений функции, $\max(y) - \min(y)$ – диапазон изменения тестовой функции.

Ниже приведены результаты вычисления MSPE для различных значений входных параметров для функции $10 \sin(\frac{2\pi x}{100})$:

Таблица 1 – Значения MSPE (Mean Squared Percentage Error) для функций $10 \sin(\frac{2\pi x}{100})$

σ шума, %	Значение ρ	MSPE
$h = 5$		
5	0,5	0,5379
5	0,8	0,3755
10	0,5	0,794
10	0,8	0,7375
$h = 10$		
5	0,5	1,2243
5	0,8	0,7799

σ шума, %	Значение ρ	MSPE
$h = 10$		
10	0,5	1,1618
10	0,8	0,7016
$h = 15$		
5	0,5	2,2775
5	0,8	1,0442
10	0,5	2,2798
10	0,8	0,9403

Таблица 2 – Значения MSPE (Mean Squared Percentage Error) для функции

$$\sin\left(\frac{\pi x}{20}\right)e^{(0,02x)} + 3$$

σ шума, %	Значение ρ	MSPE
$h = 5$		
5	0,5	0,7586
5	0,8	0,7801
10	0,5	0,9961
10	0,8	1,3646
$h = 10$		
5	0,5	4,1631
5	0,8	2,937
10	0,5	4,2798
10	0,8	2,8226
$h = 15$		
5	0,5	7,4581
5	0,8	4,5543
10	0,5	7,175
10	0,8	4,7624

Как можно заметить, в зависимости от значений параметров сплайна меняется значение оценки качества, но при этом набор наиболее оптимальных параметров зависит также от того, какой процент зашумления полезного сигнала применялся к функции.

При применении штрафного сплайна непосредственно к задаче прогнозирования на основе пользовательских данных были использованы следующие наборы данных:

- значения цены закрытия акций компании «Мегафон» в срок с 01.06.2017 до 01.11.2017;

- значения цены закрытия акций банка «Сбербанк» в срок с 01.11.2017 до 26.02.2018. [32]

Графики изменения значений цен изображены на рисунках 16-17.

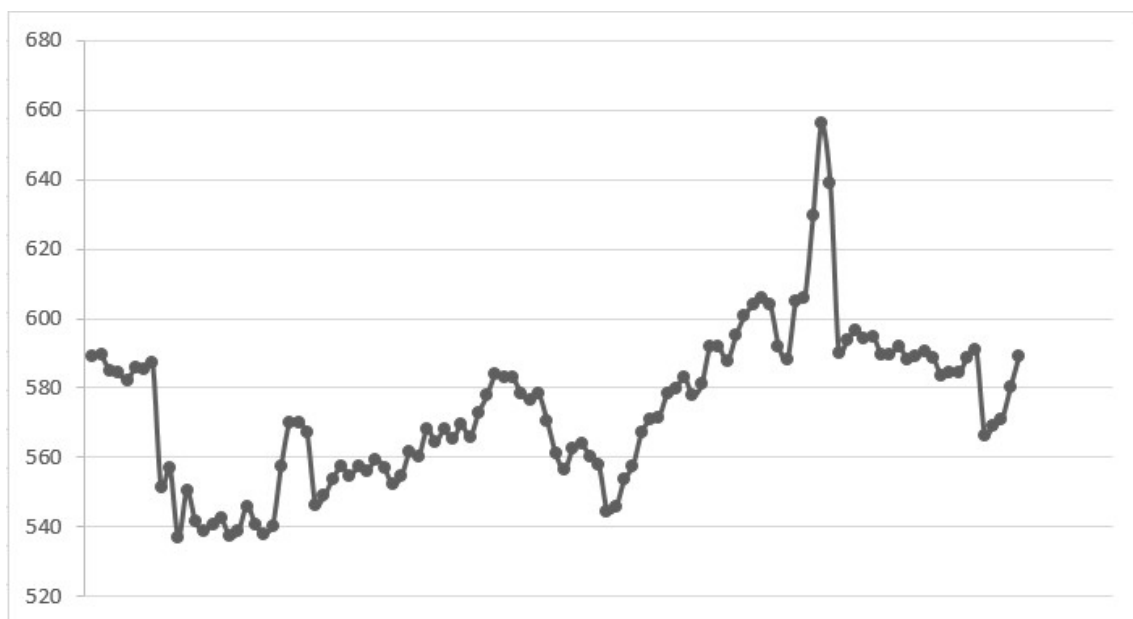


Рисунок 16 – График изменения значений цен на акции компании «Мегафон»

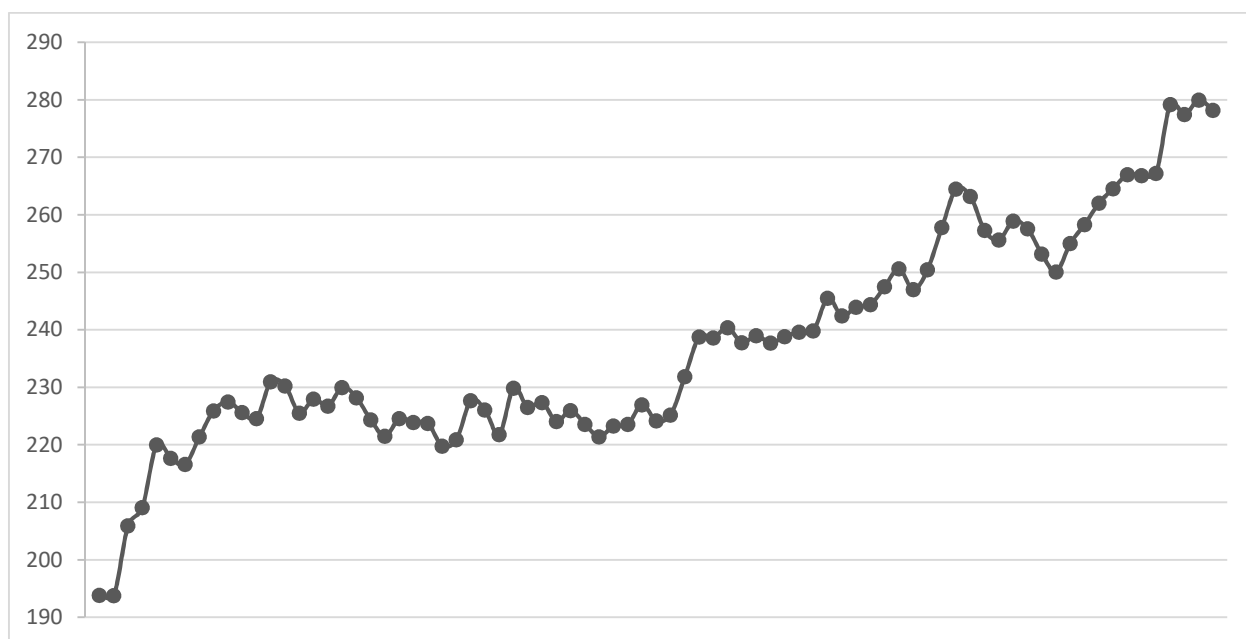


Рисунок 17 – График изменения значений цен на акции банка «Сбербанк»

Значения среднеквадратичной погрешности прогнозирования для значений акций «Мегафон» для различных значений параметров представлены в таблице 3.

Таблица 3 – Значения MSPE (Mean Squared Percentage Error)

h = 10		h = 15		h = 20	
ρ	σ	ρ	σ	ρ	σ
0,2	11,8841	0,2	12,7737	0,2	13,7671
0,5	10,9013	0,5	11,6107	0,5	12,1041
0,8	10,6250	0,8	10,9280	0,8	11,5479

Исходя из полученных значений, также можно обратить внимание на то, что значения погрешностей меняются в зависимости от набора входных параметров сплайна.

5.2 Анализ результатов качества прогнозирования с использованием оптимизационного алгоритма

Как было отмечено ранее значение оценки качества прогнозирования зависит также от регулируемых входных параметров сплайна. Для определения оптимальных параметров в приложении используется генетический алгоритм, определяющий такие значения входных параметров, при которых погрешность прогнозных данных минимальна. Для анализа работы алгоритма оптимизации в качестве входных данных будем использовать значения цен на акции, изображенные на рисунках 14-15, а также для сравнения качества прогнозирования с использованием алгоритма оптимизации и без определим прогнозные данные для других случайных наборов регулируемых параметров. Для представления результата исследования было принято решение использовать диаграмму «ящик с усами» или диаграмму размаха (box-and-whiskers plot), так как данный тип диаграмм компактен и очень прост для понимания.

Один ящик содержит выборку из всех значений для определенного набора параметров. Линия внутри каждого ящика показывает значение медианы, нижняя и верхняя грань соответствует первому и третьему квартилю

(значениям, отделяющим $\frac{1}{4}$ и $\frac{3}{4}$ выборки), горизонтальные черточки на конце «усов» - максимальное и минимальное значение выборки. Следует учесть, что в ящик не включаются выбросы – значения, выходящие за 3 среднеквадратичных отклонения, поэтому максимум и минимум выборки отображается без их участия.

На рисунке 18 изображены ящики с усами для значений цены закрытия акций компании «Мегафон» в срок с 01.06.2017 до 01.11.2017. Первые пять ящичков построены из случайно подобранных значений входных параметров, последний столбец построен на основе значений параметров $h = 4$ и $\rho = 0,87$. Как видно, из графика данная комбинация имеет наименьший разброс погрешностей, по сравнению с остальными и поэтому, можно предположить, что является наиболее оптимальной для данного набора данных. Штрафной сплайн, построенный с использованием данных входных параметров изображен на рисунке 19, значения MSPE для каждого ящика приведены в таблице 4.

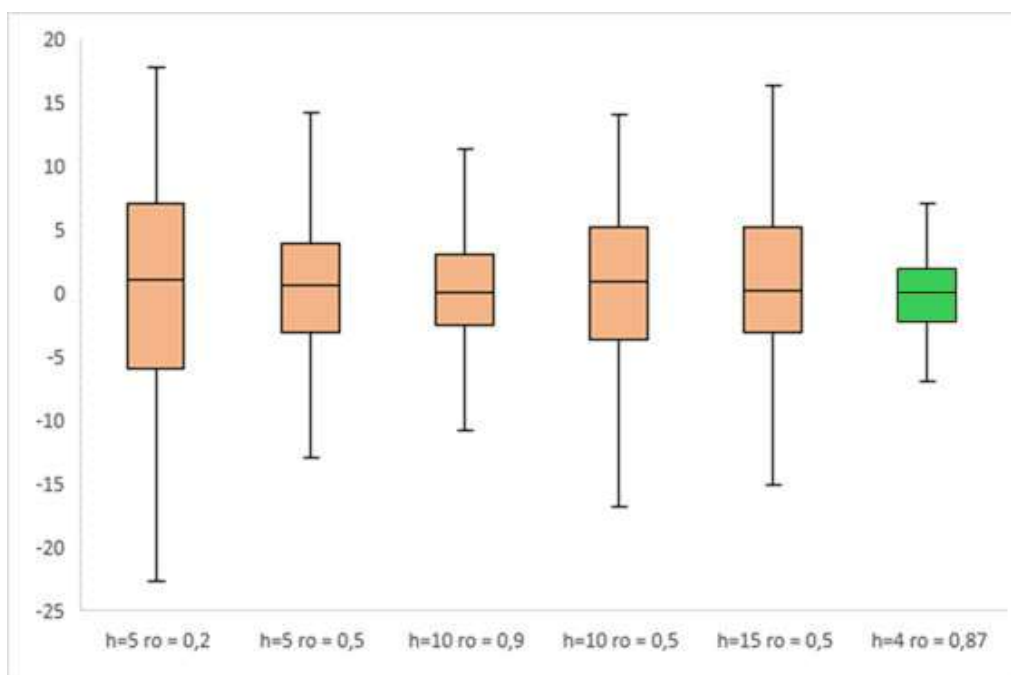


Рисунок 18 – Сравнение разбросов погрешностей для различных входных параметров сплайна

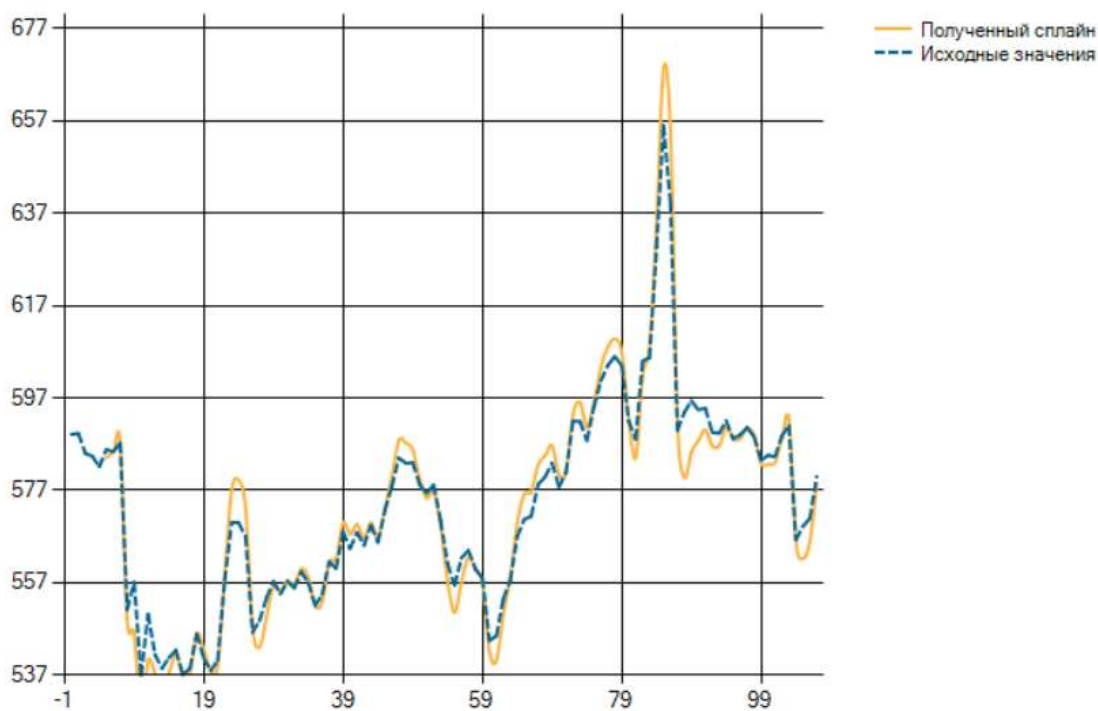


Рисунок 19 – График полученного сплайна при оптимальных параметрах

Таблица 4 – Значения MSPE для различных наборов параметров

Параметры	$h=5, \rho=0,2$	$h=5, \rho=0,5$	$h=10, \rho=0,5$	$h=10, \rho=0,9$	$h=15, \rho=0,5$	$h=4, \rho=0,87$
MSPE. %	8.575	5.564	6.550	5.218	7.535	3.643

Аналогичный поиск оптимальных параметров и сравнение его с случайно подобранными параметрами был произведен для другого набора данных – значения цены закрытия акций банка «Сбербанк» в срок с 01.11.2017 до 26.02.2018 (рисунок 20). В данном случае, значения оптимальных параметров, полученных в результате применения алгоритма оптимизации – $h = 7$ и $\rho = 0,91$. При этом, на графике видно, что разброс погрешностей для данных значений параметров больше, чем значения в соседнем графике с параметрами $h = 5$ и $\rho = 0,9$. Это объясняется тем, что используемая в оценке качества прогнозирования процентная погрешность, как было отмечено при анализе оценок прогнозирования, чувствительна к выбросам, и поэтому реагирует на выбросы, присутствующие в наборе погрешностей с параметрами $h = 5$ и $\rho = 0,9$. На графике «ящик с усами» выбросы обычно не отображаются. Значения процентной погрешности для приведенных наборов

параметров приведены в таблице 5. Штрафной сплайн, построенный с использованием данных входных параметров изображен на рисунке 21.

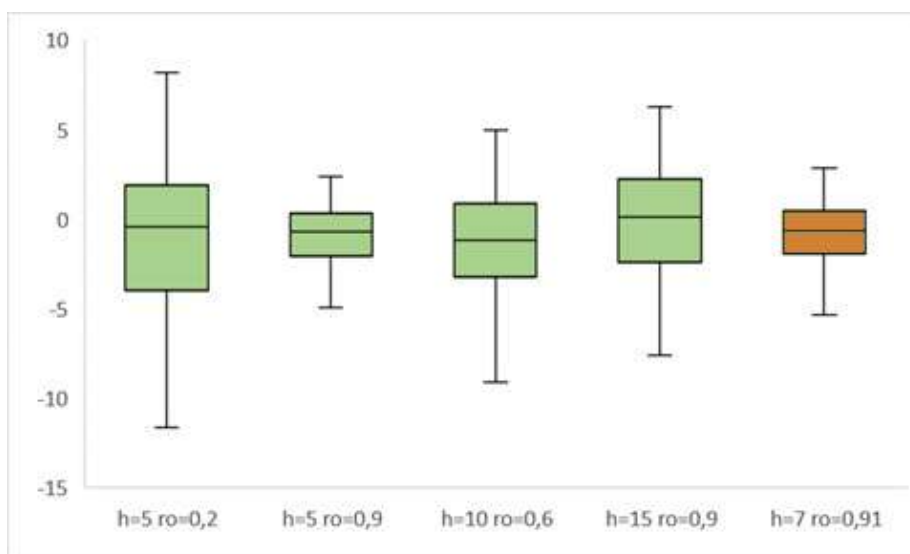


Рисунок 20 – Сравнение разбросов погрешностей для различных входных параметров сплайна

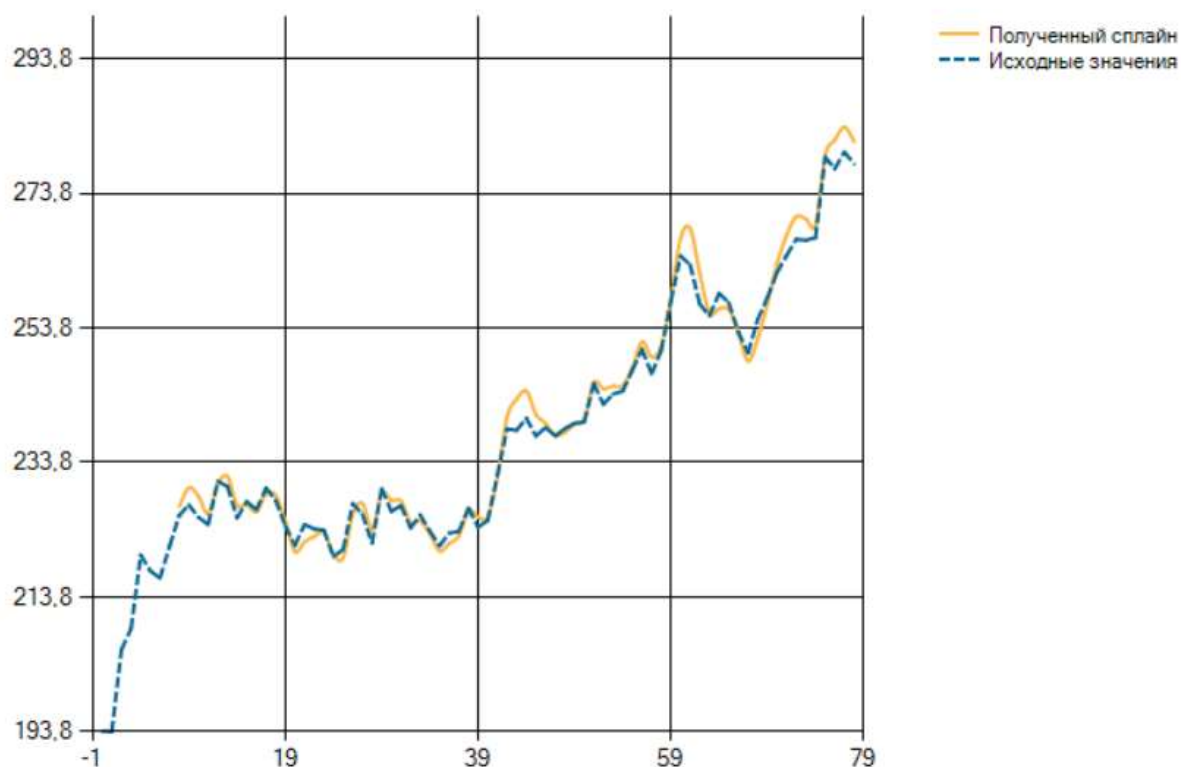


Рисунок 21 – График полученного сплайна при оптимальных параметрах

Таблица 5 – Значения MSPE для различных наборов параметров

Параметры	$h=5, \rho=0,2$	$h=5, \rho=0,9$	$h=10, \rho=0,6$	$h=15, \rho=0,9$	$h=7, \rho=0,91$
MSPE. %	3,210	3,101	5,558	5,239	2,151

Исходя из значений оценок качества прогнозирования, погрешность прогнозирования при использовании штрафных сплайнов с определением оптимальных параметров находится в пределах значений от 2 до 4 %, что является довольно приемлемым значением при сравнении с другими методами прогнозирования, рассмотренными ранее. При этом, некоторые комбинации методов могут давать более точный результат (так комбинирование нейронных сетей и регрессионных методов, например ARIMA, дает значение ошибки около 1 %), но такие комбинации методов требуют предварительного анализа поступающих данных, что делает невозможным их работу в масштабе реального времени. Метод основанный на штрафных сплайнах обеспечивает приемлемое соотношение качества прогнозирования и скорости прогнозирования, при этом тип и природа входных данных не имеет значения и не требует предварительной обработки.

6. ФИНАНСОВЫЙ МЕНЕДЖМЕНТ, РЕСУРСОЭФФЕКТИВНОСТЬ И РЕСУРСОСБЕРЕЖЕНИЕ

На сегодняшний день научные исследования, прежде всего, оцениваются с точки зрения коммерческой ценности проведенной разработки, которая является необходимым условием, например, для поиска источников инвестиций и финансирования научного исследования, а также для коммерциализации его результатов. Основными вопросами для определения коммерческой ценности являются вопрос востребованности продукта на рынке, определение его цены и стоимости научного проекта, сроки, которые необходимы для того, чтобы выпустить продукт на рынок.

Следовательно, целью раздела «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение» является обеспечение создания и проектирования такой разработки или технологии, которая будет конкурентоспособна на рынке, а также отвечать современным требованиям в области ресурсоэффективности и ресурсосбережения.

Для достижения этой цели в данной работе решаются следующие задачи:

- в первой части раздела «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение» оценивается коммерческий потенциал, перспективность научного исследования и возможные альтернативы с позиции ресурсоэффективности и ресурсосбережения;
- вторая часть раздела содержит планирование научного исследования и формирование его бюджета;
- в заключительной части определяется ресурсная, финансовая, социальная и экономическая модель исследования.

6.1. Оценка коммерческого потенциала, перспективности и альтернатив проведения НИ с позиции ресурсоэффективности и ресурсосбережения

6.1.1. Потенциальные потребители результатов исследования

Для анализа потребителей описываемого в рамках данной работы исследования необходимо определить целевой рынок разработки и провести его дальнейшее сегментирование.

Целевым рынком разработки является рынок систем и моделей прогнозирования временных рядов. В качестве первого критерия сегментирования выделим вид предоставляемого для прогнозирования ресурса. Наиболее популярными на данный момент являются библиотеки кода, представляющие собой сборники подпрограмм и объектов для их дальнейшего использования в разработке и исследованиях, и модули, предназначенные для запуска и дальнейшего использования внутри математических пакетов (MatLab, MathCad и т.д.). Кроме этого, используются в качестве ресурсов веб-приложения и десктоп-программы. Вторым критерием можно выделить профессиональные навыки конечного пользователя. Некоторые типы разработок требуют обязательного наличия навыка программирования, другие – навыка работы со специальными математическими приложениям, также наличие специальных профессиональных навыков может и не требоваться.

На основании выделенных выше критериев построена и приведена на таблице 6 карта сегментирования рынка. Темным цветом выделены наиболее популярные на данный момент решения. Светлым – существующие, но менее распространенные решения. Решения, выделенные белым – не используются или их использование не имеет смысла.

Таблица 6 – Карта сегментирования рынка систем и моделей прогнозирования

	Тип предоставляемого ресурса				
		Десктоп-приложение	Библиотека кода	Модуль математического пакета	Веб-приложение
Профессиональные навыки конечного пользователя	Навыки программирования				
	Навыки работы с математическими пакетами				
	Отсутствие особых навыков				

Анализируя полученную карту сегментирования, можно прийти к выводу, что на данный момент наиболее перспективны в использовании десктоп- и веб-приложения, так как они не требуют наличия специальных навыков у конечного пользователя, а также достаточно небольшое количество конкурентных решений в данных видах, позволяют выбрать один из них в качестве наиболее привлекательного сегмента для направления разработки.

6.1.2. Анализ конкурентных технических решений

Выделим некоторые существующие аналоги разрабатываемой системы:

- Библиотека кода Forecast для языка программирования R;
- Математический пакет Neural Networks Toolbox для работы с нейронными сетями в среде MatLab
- Программный продукт Novo Forecast Pro

Чтобы произвести качественную оценку сравнительной эффективности научной разработки и определить направления будущего развития, необходимо проанализировать приведенные выше конкурентные решения. Результаты данного анализа приведены в таблице 7.

Таблица 7 – Оценочная карта для сравнения конкурентных технических решений

Критерии оценки	Вес критерия	Баллы				Конкурентоспособность			
		Бф	Бк1	Бк2	Бк3	Кф	Кк1	Кк2	Кк3
Технические критерии оценки ресурсоэффективности									
1. Простота эксплуатации	0,2	5	2	1	5	0,2	0,08	0,04	0,2
2. Понятность интерфейса	0,15	5	2	2	5	0,15	0,03	0,03	0,15
3. Функциональные возможности	0,2	3	5	3	5	0,12	0,2	0,12	0,2
4. Скорость внедрения	0,05	5	5	4	3	0,05	0,05	0,04	0,03
5. Потребность в ресурсах памяти	0,025	4	4	4	2	0,02	0,02	0,02	0,01
Экономические критерии оценки ресурсоэффективности									
1. Цена	0,15	4	5	4	2	0,12	0,15	0,12	0,06
2. Область применения	0,15	4	5	3	5	0,12	0,15	0,09	0,15
3. Наличие обновлений	0,025	5	5	5	5	0,025	0,025	0,025	0,025
4. Уровень проникновения на рынок	0,05	0	5	4	4	0	0,05	0,04	0,04
Итого:	1					0,805	0,675	0,525	0,865

В результате анализа конкурентных разработок можно прийти к выводу, что разрабатываемое программное обеспечение является конкурентоспособным и может быть представлено на рынке. Преимущества перед конкурентами у разрабатываемого приложения особенно отчетливо проявляется в таких параметрах как простота эксплуатации и понятность интерфейса, что очень высоко оценивается из-за увеличения числа потенциальных пользователей системы. Разрабатываемое программное приложение уступает конкурентам в широте функционала, так что при дальнейшей модернизации приложения стоит обратить на данный критерий особенное внимание.

6.1.3. Технология QuaD

Технология QuaD представляет собой инструмент измерения характеристик, которые описывают качество научного исследования и разработки, а также перспективность разработки на рынке, и предназначены для принятия решения целесообразности вложения денежных средств в научно-исследовательский проект.

В соответствии с технологией QuaD по стобалльной шкале оценивается каждый показатель из группы показателей оценки коммерческого потенциала разработки и группы показателей качества разработки. Веса показателей, определяемые экспертным путем, в сумме должны составлять 1.

Оценка качества и перспективности разработки по технологии QuaD представлена в таблице 8.

Таблица 8 – Оценка качества и перспективности разработки по технологии QuaD

Критерии оценки	Вес критерия	Баллы	Максимальный балл	Относительное значение	Средневзвешенное значение
Показатели оценки качества разработки					
1. Унифицированность	0,075	85	100	0,85	6,375
2. Уровень материалоемкости разработки	0,075	90	100	0,9	6,75
3. Функциональные возможности	0,2	75	100	0,75	15
4. Простота эксплуатации	0,2	100	100	1	20
5. Устойчивость	0,05	95	100	0,95	4,75
Показатели коммерческого потенциала разработки					
1. Конкурентоспособность разработки	0,15	75	100	0,75	11,25
2. Перспективность рынка	0,15	90	100	0,9	13,5
3. Финансовая эффективность научной разработки	0,075	90	100	0,9	6,75
4. Поддержка продукта	0,025	40	100	0,4	1
Итого	1				85,375

Полученное результирующее значение позволяет говорить о перспективах разработки и качестве проведенного исследования. Значение находится в диапазоне от 80 до 100, а значит такая разработка считается перспективной. По результатам оценки качества и перспективности делается вывод об объемах инвестирования в текущую разработки и направлениях ее дальнейшей модернизации.

6.1.4. SWOT-анализ

SWOT-анализ представляет собой комплексный анализ научного исследования. Прежде всего, на первом этапе SWOT-анализа для проведения анализа необходимо определить сильные и слабые стороны научного проекта, а также угрозы и возможности, которое могут проявиться или уже проявились во внешней среде.

Сильные стороны научного исследования характеризуют те преимущества проекта, которые являются особенными с точки зрения конкуренции, то есть ресурсы и возможности, которые могут быть для достижения поставленных целей разработки данного проекта. Слабые стороны, наоборот, характеризуют те особенности проекта, которые препятствуют достижению целей. Возможности представляют собой любую предположительную ситуацию, которая может возникнуть в настоящем или будущем, в условиях внешней среды проекта, и, при этом, улучшить спрос на результаты проекта или конкурентную позицию разработки на рынке. Угрозы – это также возможные в настоящем или будущем ситуации во внешней среде, но влияющие отрицательно на научный проект, угрожающие его конкурентоспособности и спросу.

Результаты первого этапа SWOT-анализа представлены в таблице 9.

Таблица 9 – Промежуточная матрица SWOT

	Положительное влияние	Отрицательное влияние
Внутренняя среда предприятия	<i>Сильные стороны научного исследования:</i> С1. Понятный и «дружественный» пользовательский интерфейс С2. Простота эксплуатации С3. Небольшие затраты ресурсов ОС	<i>Слабые стороны научного исследования:</i> Сл1. Ограниченность функциональных возможностей системы Сл2. Невозможность пользовательской настройки интерфейса Сл3. Неизвестность используемой модели прогнозирования на целевом рынке
Внешняя среда предприятия	<i>Возможности научного исследования:</i> В1. Заинтересованность в разработке смежных сегментов рынка В2. Использование разработки в образовательных целях В3. Увеличение популярности прогнозирования в различных сферах производства	<i>Угрозы научного исследования:</i> У1. Появление систем-конкурентов У2. Появление новых моделей прогнозирования с более высоким уровнем точности У3. Потеря интереса клиентов к десктоп-приложениям

После того, как для всех областей SWOT-анализа сформулированы пункты, на втором этапе выявляются соответствия сильных и слабых сторон научного проекте внешним условиям окружающей среды. Это необходимо для выявления степени необходимости проведения стратегических изменений.

Результатами второго этапа являются интерактивные матрицы проекта, представленные в таблицах 10-13.

Таблица 10 – Интерактивная матрица проекта

Сильные стороны научного исследования				
		C1	C2	C3
Возможности научного исследования	B1	+	+	0
	B2	+	+	+
	B3	+	+	0

Таблица 11 – Интерактивная матрица проекта

Сильные стороны научного исследования				
Угрозы научного исследования		С1	С2	С3
	У1	+	+	0
	У2	-	-	-
	У3	+	0	-

Таблица 12 – Интерактивная матрица проекта

Слабые стороны научного исследования				
Возможности научного исследования		Сл1	Сл2	Сл3
	В1	+	+	0
	В2	0	0	+
	В3	-	+	+

Таблица 13 – Интерактивная матрица проекта

Слабые стороны научного исследования				
Угрозы научного исследования		Сл1	Сл2	Сл3
	У1	+	+	+
	У2	-	-	0
	У3	-	-	-

В рамках третьего этапа должна быть составлена итоговая матрица SWOT-анализа. Конечные результаты учитываются при разработке структуры работ, выполняемых в рамках научного исследования. Итоговая матрица представлена в приложении В.

6.1.5. Определение возможных альтернатив проведения научных исследований

Для определения возможных альтернатив проведения научных исследований рекомендуется использовать морфологический подход, который основан на систематическом исследовании всех возможных в теории вариантов, возникающих из морфологии объекта исследования.

Исходя из этапов реализации метода, прежде всего необходимо определить все важные морфологические характеристики объекта исследования. К ним можно отнести:

- язык программирования;
- графический интерфейс пользователя;
- вид визуализации;

- формат сохранения данных для анализа
- способ оптимизации параметров модели прогнозирования

Далее следует раскрыть все возможные варианты для каждой характеристики. В рамках этого этапа составляется морфологическая матрица. В таблице 14 представлена морфологическая матрица проекта.

Таблица 14 – Морфологическая матрица проекта

	1	2	3
А. Язык программирования	C#	Python	Java
Б. Графический интерфейс пользователя	Десктоп-приложение	Веб-приложение	
В. Вид визуализации	График «Ящик с усами»	График «Слайн»	Табличный вид
Г. Формат сохранения данных для анализа	Формат Excel-документа	Формат .txt	Формат .csv
Д. Способ оптимизации параметров модели прогнозирования	Нейронные сети	Метод Хука-Дживса	Генетический алгоритм

Из данной морфологической матрицы можно выделить три варианта реализации проекта:

Исполнение 1: А1Б1В2Г1Д3

Исполнение 2: А2Б2В2С2Д1

Исполнение 3: А3А1В1Г1Д2

6.2. Планирование и формирование бюджета научных исследований

6.2.1. Структура работ в рамках научного исследования

При планировании комплекса предполагаемых работ был разработан перечень этапов и работ в рамках проведения научно-исследовательского проекта. В данном проекте исполнителями являются студент и научный руководитель. Перечень этапов, работ и распределение исполнителей по данным видам работ в рамках проводимого научно-исследовательского проекта представлен в приложении Г.

6.2.2. Определение трудоемкости выполнения работ

Для последующего определения стоимости разработки прежде всего необходимо определить трудовые затраты участников научно-

исследовательского проекта, поэтому необходимо определить трудоемкость работ каждого из участников научного исследования.

Для определения ожидаемого (среднего) значения трудоемкости используется следующая формула:

$$t_{ожi} = \frac{3t_{\min i} + 2t_{\max i}}{5},$$

где $t_{ожi}$ – ожидаемая трудоемкость выполнения i -ой работы в человеко-днях, $t_{\min i}$ – минимально возможная трудоемкость выполнения заданной i -ой работы в человеко-днях, а $t_{\max i}$ – максимально возможная трудоемкость выполнения заданной i -ой работы в человеко-днях.

На основе ожидаемой трудоемкости работ, можно определить продолжительность каждой работы в рабочих днях, учитывая, при этом, параллельность выполнения работ несколькими исполнителями. Для этого используется следующая формула:

$$T_{pi} = \frac{t_{ожi}}{Ч_i},$$

где T_{pi} – продолжительность одной работы в рабочих днях, $Ч_i$ – численность исполнителей, выполняющих одновременно одну и ту же работу на данном этапе.

Для удобства построения графика, длительность каждого из этапов работ из рабочих дней следует перевести в календарные дни. Для этого необходимо воспользоваться следующей формулой:

$$T_{ki} = T_{pi} \cdot k_{кал},$$

где T_{ki} – продолжительность выполнения i -ой работы в календарных днях, T_{pi} – продолжительность выполнения i -ой работы в рабочих днях, $k_{кал}$ – коэффициент календарности.

Коэффициент календарности определяется по следующей формуле:

$$k_{\text{кал}} = \frac{T_{\text{кал}}}{T_{\text{кал}} - T_{\text{вых}} - T_{\text{пр}}},$$

где $T_{\text{кал}}$ – количество календарных дней в году, $T_{\text{вых}}$ – количество выходных дней в году, $T_{\text{пр}}$ – количество праздничных дней в году.

Коэффициент календарности в 2018 году – 1,477

Все расчеты по данным формулам представлены в приложении Д.

6.2.3. Разработка графика проведения научного исследования

Наиболее удобным и наглядным для представления графика проведения научного исследования является горизонтальный ленточный график диаграмма Ганта. На данной диаграмме работы по теме научно-исследовательского проекта представляются протяженными во времени отрезками с датами начала и окончания выполнения данных этапов проекта.


Диаграмма Ганта с разбивкой по месяцам и декадам за период времени дипломирования представлена в таблице 15.

Таблица 15 – Календарный план-график проведения работ

№	Вид работ	Исполнители	T_{ki}	Продолжительность выполнения работ														
				январь		февраль			март			апрель			май			
				2	3	1	2	3	1	2	3	1	2	3	1	2		
1	Разработка технического задания	Научный руководитель	1	■														
2	Выбор направления исследований	Научный руководитель, студент	16	■	■													
3	Теоретические и экспериментальные исследования	Студент	77															
4	Обобщение и оценка результатов	Научный руководитель, студент	7															
5	Разработка технической документации и проектирование	Студент	10															

№	Вид работ	Исполните ли	T _{кi}	Продолжительность выполнения работ												
				январь		февраль			март			апрель		май		
				1	2	1	2	3	1	2	3	1	2	2	1	2
6	Оформление отчета	Студент	6													

 – студент

 – научный руководитель

6.2.4. Бюджет научно-технического исследования

При планировании научно-исследовательского проекта необходимо отобразить все виды расходов, которые связаны с его выполнением. В процессе формирования затраты группируются по следующим статьям:

- материальные затраты НТИ;
- затраты на специальное оборудование для научных работ;
- дополнительная заработная плата исполнителей темы;
- отчисления во внебюджетные фонды;
- затраты на научные и производственные командировки;
- контрагентные расходы;
- накладные расходы.

6.2.4.1. Расчет материальных затрат научно-технического исследования

Прежде всего необходимо рассчитать материальные затраты на научно-исследовательский проект. В данную статью входят затраты на сырье и материалы, покупные материалы, используемые в процессе создания научно-технической продукции и т.д. В материальные затраты, помимо вышеуказанных, включаются дополнительно затраты на канцелярские принадлежности, диски, картриджи и т.п.

Расчет материальных затрат осуществляется по следующей формуле:

$$Z_{.m} = (1 + k_T) \sum_{i=1}^m C_i N_{расхi},$$

где m – количество материальных ресурсов i -го вида, планируемых к использованию при выполнении научного исследования, C_i – цена приобретения единицы i -го вида потребляемых ресурсов, k_T – коэффициент, учитывающий транспортно-заготовительные расходы.

Материальные затраты, необходимые для данной разработки, занесены в таблицу 16.

Таблица 16 – Материальные затраты

Наименование	Единица измерения	Количество			Цена за ед., руб.			Затраты на материалы, руб		
		И1	И2	И3	И1	И2	И3	И1	И2	И3
Бумага, формат А4	Лист	150	150	150	0,5	0,5	0,5	75	75	75
Брюшорование	Штука	1	1	1	100	100	100	100	100	100
Итого								175	175	175

6.2.4.2. Основная заработная плата исполнителей темы

В данную статью бюджета включается основная заработная плата научных работников, непосредственно участвующих в выполнении работ по данной теме. Величина расходов по данной статье определяется исходя из трудоемкости выполняемых работ и действующей системы окладов и тарифных ставок.

Статья включает основную заработную плату работников, непосредственно занятых выполнением научно-техническим исследованием (включая премии и доплаты) и дополнительную заработную плату (12-20 % от основной заработной платы):

$$Z_{zn} = Z_{осн} + Z_{доп}.$$

Основная заработная плата рассчитывается по следующей формуле:

$$Z_{осн} = Z_{дн} T_p,$$

где T_p – продолжительность работ, выполняемых научно-техническим работником, $Z_{дн}$ – среднедневная заработная плата работника, которая рассчитывается по формуле:

$$Z_{\text{дн}} = \frac{Z_m M}{F_d},$$

где Z_m – месячный должностной оклад работника, M – количество месяцев работы без отпуска в течение года, F_d – действительный годовой фонд рабочего времени научно-технического персонала. Баланс рабочего времени приведен в таблице 17.

Таблица 17 – Баланс рабочего времени

Показатели рабочего времени	Руководитель	Студент
Календарное число дней	365	365
Количество нерабочих дней - выходные дни - праздничные дни	118	118
Потери рабочего времени - Отпуск - Невыходы по болезни	24	24
Действительный годовой фонд рабочего времени	223	223

Месячный должностной оклад работника:

$$Z_m = Z_{mc} (1 + k_{np} + k_d) k_p,$$

где Z_{mc} – заработная плата по тарифной ставке, руб; k_{np} – премиальный коэффициент (равный 30 % от Z_{mc}), k_d – коэффициент доплат и надбавок (примерно 0,2-0,5), k_p – районный коэффициент (для Томска – 1,3).

Таблица 18 – Расчет основной заработной платы

Исполнители	Разряд	Оклад	k_{np}	k_p	Z_m	$Z_{дн}$	T_p	$Z_{осн}$
Руководитель	4-ый	16819,03	0,3	0,5	39356,53	1293,91	4,8	6210,78
Студент	1-ый	9893,56	0,3	0,2	19291,35	634,23	70,6	44777,07

Размер основной заработной платы для других исполнений приведен в таблице 19.

6.2.4.3. Дополнительная заработная плата исполнителей темы

Расчет дополнительной заработной платы ведется по следующей формуле:

$$Z_{\text{доп}} = k_{\text{доп}} Z_{\text{осн}},$$

где $k_{доп}$ – коэффициент дополнительной заработной платы (на стадии проектирования принимается равным 0,12 – 0,15).

6.2.4.4. Отчисления во внебюджетные фонды (страховые отчисления)

В данной статье расходов отражаются обязательные отчисления по установленным законодательством Российской Федерации нормам органам государственного социального страхования (ФСС), пенсионного фонда (ПФ) и медицинского страхования (ФФОМС) от затрат на оплату труда работников.

Величина отчислений во внебюджетные фонды определяется исходя из следующей формулы:

$$З_{внеб} = k_{внеб} (З_{осн} + З_{доп}),$$

где $k_{внеб}$ – коэффициент отчислений на уплату во внебюджетные фонды (пенсионный фонд, фонд обязательного медицинского страхования и пр.). Для образовательных учреждений это значение равно 27,1 %.

Отчисления во внебюджетные фонды представлены в таблице 19.

Таблица 19 – Отчисления во внебюджетные фонды

Исполнитель	Основная заработная плата			Дополнительная заработная плата		
	И1	И2	И3	И1	И2	И3
Руководитель проекта	6210,78	6210,78	6210,78	745,29	745,29	745,29
Студент-дипломник	44777,07	50738,4	50738,4	5373,25	6088,61	6088,61
Коэффициент отчислений во внебюджетные фонды			27,1 %			
Итого						
Исполнение 1	72582,05					
Исполнение 2	81068,2					
Исполнение 3	81068,2					

6.2.4.5. Накладные расходы

Накладные расходы учитывают прочие затраты организации, не попавшие в предыдущие статьи расходов: печать и ксерокопирование материалов исследования, оплата услуг связи, электроэнергии, почтовые и телеграфные расходы, размножение материалов и т.д. Их величина определяется по следующей формуле:

$$Z_{\text{накл}} (\text{сумма}_{\text{ статей}} / \text{кол-во}_{\text{ статей}}) * k_{\text{пр}}$$

где $k_{\text{пр}}$ – коэффициент, учитывающий накладные расходы.

Величину коэффициента накладных расходов можно взять в размере 16%.

6.2.4.6. Формирование бюджета затрат научно-исследовательского проекта

Определение бюджета затрат на научно-исследовательский проект по каждому варианту исполнения приведено в таблице 17.

Таблица 20 – Расчет бюджет затрат НИИ

Наименование статьи	Сумма, руб		
	И1	И2	И3
1. Материальные затраты НИИ	175	175	175
2. Затраты на специальное оборудование для научных работ	0	0	0
3. Затраты по основной заработной плате исполнителей темы	50987,85	56949,18	56949,18
4. Затраты по дополнительной заработной плате исполнителей темы	6118,54	6833,9	6833,9
5. Отчисления во внебюджетные фонды	15475,74	17285,12	17285,12
6. Затраты на научные и производственные командировки	0	0	0
7. Контрагентские расходы	0	0	0
8. Накладные расходы	2690,01	3249,728	3249,728
9. Бюджет затрат НИИ	69940,14	84492,7	84492,7

6.3. Определение ресурсной (ресурсосберегающей), финансовой, бюджетной, социальной и экономической эффективности исследования

Для определения эффективности используется расчет интегральных показателей эффективности научно-исследовательского проекта.

Интегральный показатель финансовой эффективности получают при оценке бюджета затрат нескольких вариантов использования. Для его расчета наибольший интегральный показатель реализации технической задачи принимается за базу расчета, с ним соотносятся финансовые значения по остальным вариантам исполнения.

Интегральный финансовый показатель разработки можно определить по формуле:

$$I_{финр}^{испi} = \frac{\Phi_{pi}}{\Phi_{max}},$$

где $I_{финр}^{испi}$ – интегральный финансовый показатель разработки, Φ_{pi} – стоимость i -го варианта исполнения, Φ_{max} – максимальная стоимость исполнения научно-исследовательского проекта.

Интегральный показатель ресурсоэффективности вариантов исполнения объекта можно определить по следующей формуле:

$$I_{pi} = \sum a_i b_i,$$

Где I_{pi} – интегральный показатель ресурсоэффективности для i -го варианта исполнения разработки; a_i – весовой коэффициент i -го варианта исполнения разработки; b_i – бальная оценка i -го варианта исполнения разработки, устанавливается экспертным путем по выбранной шкале оценивания.

Таблица 21 – Сравнительная оценка характеристик вариантов исполнения проекта

Критерии	Весовой коэффициент параметра	И1	И2	И3
1. Способствует росту производительности труда пользователя	0,1	5	5	5
2. Удобство в эксплуатации (соответствует требованиям потребителя)	0,15	5	5	3
3. Помехоустойчивость	0,15	4	3	4
4. Энергосбережение	0,20	4	4	4
5. Надежность	0,25	5	4	5
6. Материалоемкость	0,15	4	4	4
Итого:	1	4,5	4,1	4,2

Интегральный показатель эффективности вариантов использования разработки определяется на основании интегрального показателя по формуле:

$$I_{исп1} = \frac{I_{p-исп1}}{I_{финр}^{исп1}}, I_{исп2} = \frac{I_{p-исп2}}{I_{финр}^{исп2}} \text{ и т.д.}$$

Сравнительная эффективность проекта вычисляется как отношение интегральных показателей эффективности вариантов исполнения разработки.

Вычисления, необходимые для определения сравнительной эффективности разработки для поиска наиболее эффективного решения задачи, представленной в бакалаврской работе, приведены в таблице 22.

Таблица 22 – Сравнительная эффективность разработки

№ п/п	Показатели	Исполнение 1	Исполнение 2	Исполнение 3
1	Интегральный финансовый показатель разработки	0,828	1	1
2	Интегральный показатель ресурсоэффективности разработки	4,5	4,1	4,1
3	Интегральный показатель эффективности	5,435	4,1	4,1
4	Сравнительная эффективность вариантов исполнения	1	0,754	0,754

Исходя из результатов сравнительной эффективности разработки наиболее эффективным решением с позиции финансовой и ресурсной эффективности является вариант исполнения под номером 1.

7. СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ

Целью данной выпускной квалификационной работы является разработка программного приложения для краткосрочного прогнозирования временных рядов на основе штрафных р-сплайнов. Приложение обеспечивает пользователю возможность загрузки собственных данных для дальнейшего определения прогнозных значений, а также возможность самому ввести настраиваемые параметры сплайна или получить наиболее подходящие для данного набора данных оптимизированные значения. В качестве результата работы приложения пользователю предоставляется графическое отображение прогнозных значений и табличное представление, экспортированное в Excel.

Возможными областями применения для данного приложения являются образовательная и исследовательская деятельность, а также финансовые, промышленные, энергетические и другие предприятия, нуждающиеся в возможности краткосрочного прогнозирования в масштабе реального времени.

Так как приложение будет использоваться на рабочем месте оператора персонального компьютера, то в данном разделе рассмотрены вредные и опасные факторы, которые могут возникнуть при использовании средств вычислительной техники. Также раздел включает в себя выявление возможных вредных воздействий на окружающую среду, программ по их снижению и экономии невозполнимых ресурсов и способах защиты в чрезвычайных ситуациях, которые могут возникнуть на рабочем месте.

7.1. Производственная безопасность

Чтобы обеспечить производственную безопасность человека на рабочем месте, необходимо проанализировать возможное воздействие вредных и опасных факторов, которые могут возникнуть при разработке и использовании научно-исследовательского проекта.

Производственные факторы считаются вредными, если их воздействие на работника могут привести к его заболеванию. Производственные факторы

относятся к опасным, если их воздействие могут привести к травме работника.
[32]

Воздействие вредных факторов приводит к снижению работоспособности, появлению и развитию утомления, связанного с изменениями, возникающими в работе центральной нервной системы. При этом вредные факторы проявляются при условии интенсивности и длительности воздействия. В отличие от них, опасные факторы оказывают моментальное влияние на здоровье работника, они могут привести к травмам и ожогам различной степени тяжести, или повлиять на здоровье работника в результате отравления или облучения. В таблице 23 предоставлены возможные вредные и опасные факторы, которые могут возникнуть при работе с данным приложением.

Таблица 23 – Вредные и опасные факторы, возникающие при работе с приложением для краткосрочного прогнозирования временных рядов

Наименование видов работ	Факторы по ГОСТ 12.0.003-2015	Нормативные документы
Вредные факторы		
Работа за персональным компьютером (ПК)	Отклонение показателей микроклимата (температуры и влажности воздуха)	СанПиН 2.2.2/2.4.1340-03 [33] СанПиН 2.2.4.548-96 [34]
	Недостаточная освещенность рабочей зоны	СанПиН 2.2.2/2.4.1340-03
Опасные факторы		
Работа за персональным компьютером (ПК)	Опасность поражения электрическим током	ГОСТ 12.1.038–82 [35]
	Пожаровзрывоопасность	ФЗ от 22.07.2008 №123-ФЗ [36]

7.1.1. Вредные опасные факторы

7.1.1.1. Отклонение показателей микроклимата в помещении

Микроклимат в помещении является одним из самых необходимых для обеспечения благоприятных условий труда для работников производственным фактором, так как он оказывает большое влияние на тепловое самочувствие человека. Микроклимат в производственном помещении, прежде всего,

зависит от внешних условий таких, как категория работ, период года, условия вентиляции, а также от особенностей самого технологического процесса.

Можно выделить следующие параметры, характеризующие микроклимат в производственных помещениях:

- температура воздуха (t , °С);
- температура поверхностей (t , °С);
- относительная влажность воздуха (φ , %);
- скорость движения воздуха (v , м/с);
- интенсивность теплового облучения (I , Вт/м²).

В производственных помещениях при работе с персональными вычислительными машинами происходит постоянное выделение тепла вычислительной техникой, средствами освещения и другими вспомогательными приборами. И так как оператор ПК расположен в непосредственной близости с источниками выделения тепла, то высокая температура воздуха может способствовать быстрому перегреву программиста и быстрой утомляемости [36].

Помимо температуры воздуха, большое влияние на терморегуляцию организма оказывает относительная влажность воздуха. Как отмечено в [37] высокие показатели влажности (значения больше 85 %) затрудняют терморегуляцию тела и снижают возможность испарения пота, низкие показатели влажности (значения менее 20 %) вызывают пересыхание слизистых оболочек.

Работа программиста относится к категории Ia, то есть она производится сидя и сопровождается незначительным физическим напряжением. Интенсивность энерготрат организма для данной категории работ составляет до 120 ккал/ч (до 139 Вт).

Оптимальные микроклиматические условия обеспечивают общее и локальное ощущение теплового комфорта при минимальном напряжении механизмов терморегуляции, не вызывают отклонений в состоянии здоровья,

создают предпосылки для высокого уровня работоспособности и являются предпочтительными на рабочих местах. Оптимальные значения показателей микроклимата на рабочих местах производственных помещений согласно СанПиН 2.2.4.548-96 для категории работ Ia представлены в таблице 24. [34]

Таблица 24 – Оптимальные величины показателей микроклимата на рабочих местах производственных помещений

Период года	Категория работ	Температура воздуха, °С	Относительная влажность воздуха, %	Скорость движения воздуха, м/с
Холодный	Ia	22 – 24	40 – 60	0,1
Теплый	Ia	21 – 23	40 – 60	0,1

Допустимые микроклиматические условия не вызывают повреждений или нарушений состояния здоровья, но могут приводить к возникновению общих и локальных ощущений теплового дискомфорта, напряжению механизмов терморегуляции, ухудшению самочувствия и понижению работоспособности.

В таблице 25 приведены допустимые величины показателей микроклимата на рабочих местах производственных помещений согласно СанПиН 2.2.4.548-96 для категории работ Ia. [34]

Таблица 25 – Допустимые величины показателей микроклимата на рабочих местах производственных помещений

Период года	Категория работ	Температура воздуха, °С	Относительная влажность воздуха, %	Скорость движения воздуха, м/с
Холодный	Ia	20 – 25	15 – 75	0,1
Теплый	Ia	21 – 28	15 – 75	0,1 – 0,2

Согласно требованиям СанПиН 2.2.2/2.4.1340–03, в кабинете поддерживается температура равная 19–20 С°, при относительной влажности в 55–58%. [33] Чтобы добиться этого, необходимо проводить в помещении ежедневную влажную уборку и систематическое проветривание.

7.1.1.2. Недостаточная освещенность рабочей зоны

Недостаточная освещенность рабочей зоны – вредный производственный фактор, который регламентируется СП 52.13330.2011.

Недостаточный уровень освещенности в помещении приводит к снижению остроты зрения, головным болям, снижению концентрации внимания и, как следствие, к ухудшению производительности труда. Причиной недостаточной освещенности являются недостаточность естественного освещения, недостаточность искусственного освещения, пониженная контрастность.

В рабочем помещении должны присутствовать естественное и искусственное освещение. Коэффициент естественного освещения должен быть не менее 1,2%. Согласно СанПиН 2.2.2/2.4.1340-03 освещенность на поверхности рабочего 55 стола в зоне размещения документа должна быть 300 – 500 лк, что может достигаться установкой местного освещения, не создающего бликов на поверхности экрана. Освещенность поверхности экрана не должна превышать 300 лк. Яркость светящихся поверхностей (окон, светильников), находящихся в поле зрения должна быть не более 200 кд/м². Для источников искусственного освещения следует применять люминесцентные лампы типа ЛБ и компактные люминесцентные лампы (КЛЛ). Коэффициент пульсации при работе с компьютером не должен превышать 5%.

Помимо обеспечения достаточного уровня освещения, для минимизации данного вредного фактора следует ограничить отраженную блескость на рабочих поверхностях (экран, стол, клавиатура) за счет правильного выбора и расположения светильников, яркость бликов на экране не должна превышать 40 кд/м². Светильники местного освещения должны иметь непросвечивающий отражатель.

7.1.2. Опасные производственные факторы

7.1.2.1. Опасность поражения электрическим током

Необходимо большое внимание уделять электробезопасности при работе с электронными вычислительными устройствами, так как основная работа оператора ПК проходит за электрооборудованием.

Опасность поражения человека электрическим током оценивается величиной тока I (А), проходящего через его тело, или напряжением прикосновения U (В). Степень опасного воздействия на человека электрического тока зависит от рода и величины напряжения тока, частоты электрического тока, пути тока через тело человека, продолжительности его воздействия на организм человека, а также условий внешней среды.

Электрический ток, протекая через тело человека, производит термическое, механическое и световое воздействие – электролитическое разложение жидкости (в том числе и крови), судорожное сокращение мышц, разрыв тканей и поражение глаз.

Поражение электрическим током при работе с компьютерами возможно при наличии оголенных участков на кабеле, нарушении изоляции распределительных устройств и от токоведущих частей компьютера в случае их пробоя и нарушении изоляции, при работе с электронно-вычислительными машинами во влажной одежде и влажными руками.

Помещение, где располагается рабочее место оператора ПК, можно отнести к помещениям без повышенной опасности, так как в нем отсутствуют следующие факторы: сырость, токопроводящая пыль, токопроводящие полы, высокая температура, возможность одновременного прикосновения человека к имеющим соединение с землей металлоконструкциям зданий, технологическим аппаратам, механизмам и металлическим корпусам электрооборудования.

Для предотвращения возможности поражения электрическим током можно выделить следующие мероприятия:

- при производстве монтажных работ необходимо использовать только исправный инструмент, аттестованный службой КИПиА;
- с целью защиты от поражения электрическим током, возникающим между корпусом приборов и инструментом при пробое сетевого напряжения на корпус, корпуса приборов и инструментов должны быть заземлены;
- при включенном сетевом напряжении работы на задней панели должны быть запрещены;
- все работы по устранению неисправностей должен производить квалифицированный персонал;
- необходимо постоянно следить за исправностью электропроводки.

[33]

Согласно ГОСТ 12.1.038-82 на рабочем месте программиста допускаются уровни напряжений прикосновения и токов, представленные в таблице 26. [35]

Таблица 26 – Предельно допустимые напряжения прикосновения и токи

Род тока	Напряжение прикосновения, В	Ток, м/А
Переменный, 50 Гц	Не более 2,0	Не более 0,3
Постоянный	Не более 8,0	Не более 1,0

7.1.2.2. Пожаровзрывобезопасность

Пожарная безопасность представляет собой единый комплекс организационных, технических, режимных и эксплуатационных мероприятий по предупреждению пожаров и взрывов.

В помещениях с компьютерами повышен риск возникновения пожара из-за присутствия множества факторов: наличие большого количества электронных схем, устройств электропитания, устройств кондиционирования воздуха; возможные неисправности электрооборудования, освещения, или неправильная их эксплуатация может послужить причиной пожара.

Для устранения возможных причин возникновения пожаров необходимо проводить следующие мероприятия:

- организационные мероприятия:
 - противопожарный инструктаж обслуживающего персонала;
 - обучение персонала техники безопасности;
 - разработка инструкций, плакатов, планов эвакуации;
- эксплуатационные мероприятия:
 - соблюдение эксплуатационных норм оборудования;
 - выбор и использование современных автоматических средств тушения пожаров;
- технические мероприятия:
 - профилактический осмотр и ремонт оборудования;
 - соблюдения противопожарных мероприятий при устройстве электропроводок, оборудования, систем отопления, вентиляции и освещения

7.2. Экологическая безопасность

7.2.1. Анализ воздействия продукта на окружающую среду

Вследствие развития научно-технического прогресса постоянно увеличивается возможность воздействия на окружающую среду, создаются предпосылки для возникновения экологических кризисов.

Увеличение количества компьютерных систем, внедряемых в производственную сферу, приводит к увеличению объема потребляемой электроэнергии, что влечет за собой увеличение мощностей электростанций и их количества. И то, и другое содействует нарушению экологической обстановки и, выбросы со станций оказывают существенное влияние на атмосферу.

Основным фактором, оказывающим негативное влияние на гидросферу и литосферу, является образование отходов. В помещении образуются следующие виды отходов: бумага (макулатура), отходы от

продуктов питания и личной гигиены (упаковка, органические отходы), отходы от канцелярских принадлежностей, отходы от офисной техники (использованные картриджи, упаковка, неисправные компоненты), лампы.

7.2.2. Решения по обеспечению экологической безопасности

Наиболее активной формой защиты окружающей среды от вредного воздействия выбросов промышленных предприятий является полный переход к безотходным и малоотходным технологиям и производствам. Это потребует решения целого комплекса сложных технологических, конструкторских и организационных задач, основанных на использовании новейших научно-технических достижений.

Необходимо стремиться к снижению энергопотребления, то есть разрабатывать и внедрять системы с малым энергопотреблением. Следует использовать современные ЭВМ с режимом пониженного потребления электроэнергии при длительном простое.

Для каждого вида отходов должны применяться свои методы переработки и утилизации, недопустимо организовывать свалки мусора, проводить самостоятельно утилизацию. Все отходы следует собирать, сортировать и направлять на переработку в соответствующие организации.

В настоящее время существует ряд способов хранения и переработки твердых бытовых отходов, а именно: предварительная сортировка, санитарная земляная засыпка, сжигание, биотермическое компостирование, низкотемпературный пиролиз, высокотемпературный пиролиз.

7.3. Безопасность в чрезвычайных ситуациях

7.3.1. Перечень возможных ЧС при разработке и эксплуатации научно-исследовательского проекта

Чрезвычайные ситуации, которые могут возникнуть при разработке и эксплуатации проектируемого решения:

- техногенные (взрывы, пожары, обрушение помещений, аварии на системах жизнеобеспечения);

- природные (наводнения, ураганы, бури, природные пожары);
- биологические (эпидемии, пандемии);
- антропогенные (война, терроризм).

Общие правила поведения при чрезвычайных ситуациях:

1) Не паниковать и не поддаваться панике. Призывать окружающих к спокойствию.

2) По возможности немедленно позвонить по телефону «01», сообщить что случилось, указать точный адрес места происшествия, назвать свою фамилию и номер своего телефона.

3) Включить устройства передачи звука (радио, телевизор), а также прослушать информацию, передаваемую через уличные громкоговорители и громкоговорящие устройства. В речевом сообщении будут озвучены основные рекомендации и правила поведения.

4) Выполнять рекомендации специалистов (сотрудников полиции, медицинских работников, пожарных, спасателей).

5) Не создавать условия, которые препятствуют и затрудняют действия сотрудников полиции, медицинских работников, спасателей, пожарных.

Наиболее характерной для объекта, где размещаются рабочие помещения, оборудованные электронно-вычислительными машинами, чрезвычайной ситуацией является пожар.

Причинами возникновения данного вида ЧС могут являться:

- возникновением короткого замыкания в электропроводке;
- возгоранием устройств вычислительной техники из-за неисправности аппаратуры;
- возгоранием устройств искусственного освещения;
- возгоранием мебели по причине нарушения правил пожарной безопасности, а также неправильного использования дополнительных бытовых электроприборов и электроустановок.

Помещение для работы операторов ПЭВМ по системе классификации

категорий помещений по взрывопожарной и пожарной опасности относится к категории Д (из 5-ти категорий А, Б, В1-В4, Г, Д), т.к. относится к помещениям с негорючими веществами и материалами в холодном состоянии [39].

7.3.2. Разработка действий в результате возникшей ЧС и меры по ликвидации ее последствий

Пожарная безопасность подразумевает надлежащее состояние объекта с исключением возможности возникновения очага возгорания (пожара) и его распространения в пространстве. Обеспечение пожарной безопасности — приоритетная задача для любого предприятия. Создание системы защиты регламентировано законом и нормативными документами различных ведомств.

Каждый сотрудник организации должен быть ознакомлен с инструкцией по пожарной безопасности, пройти инструктаж по технике безопасности и строго соблюдать его.

Запрещается использовать электроприборы в условиях, не соответствующих требованиям инструкций изготовителей, или имеющие неисправности, которые в соответствии с инструкцией по эксплуатации могут привести к пожару, а также эксплуатировать электропровода и кабели с поврежденной или потерявшей защитные свойства изоляцией. Электроустановки и бытовые электроприборы в помещениях по окончании рабочего времени должны быть обесточены (вилки должны быть вынуты из розеток). Под напряжением должны оставаться дежурное освещение и пожарная сигнализация. Недопустимо хранение легковоспламеняющихся, горючих и взрывчатых веществ, использование открытого огня в помещениях офиса.

Перед уходом из служебного помещения работник обязан провести его осмотр, закрыть окна, и убедиться в том, что в помещении отсутствуют источники возможного возгорания, все электроприборы отключены и

выключено освещение. С периодичностью не реже одного раза в три года необходимо проводить замеры сопротивления изоляции токоведущих частей силового и осветительного оборудования.

Работник при обнаружении пожара или признаков горения (задымление, запах гари, повышение температуры и т.п.) должен:

- немедленно прекратить работу и вызвать пожарную охрану по телефону «01», сообщив при этом адрес, место возникновения пожара и свою фамилию;
 - принять по возможности меры по эвакуации людей и материальных ценностей;
 - отключить от сети закрепленное за ним электрооборудование; –
- Приступить к тушению пожара имеющимися средствами пожаротушения;
- сообщить непосредственному или вышестоящему начальнику и оповестить окружающих сотрудников;
 - при общем сигнале опасности покинуть здание согласно «Плану эвакуации людей при пожаре и других ЧС».

Для тушения пожара необходимо применять ручные углекислотные огнетушители (типа ОУ-2, ОУ-5), находящиеся в помещениях офиса, и пожарный кран внутреннего противопожарного водопровода. Они предназначены для тушения начальных возгораний различных веществ и материалов, за исключением веществ, горение которых происходит без доступа воздуха. Огнетушители должны постоянно содержаться в исправном состоянии и быть готовыми к действию. Категорически запрещается тушить возгорания в помещениях офиса при помощи химических пенных огнетушителей (типа ОХП-10). [40]

7.4. Правовые и организационные вопросы обеспечения безопасности

7.4.1. Специальные правовые нормы трудового законодательства

Режим труда и отдыха предусматривает соблюдение определенной длительности непрерывной работы на персональном компьютере (ПК) и перерывов, регламентированных с учетом продолжительности рабочей смены, видов и категории трудовой деятельности.

Вид трудовой деятельности на персональном компьютере в рамках данной работы соответствует группе В – творческая работа в режиме диалога с ПК, категория трудовой деятельности – III (до 6 часов непосредственной работы на ПК).

При 8-часовой рабочей смене и работе на ПК, соответствующей описанным выше критериям необходимо через 1,5- 2,0 часа от начала рабочей смены и через 1,5-2,0 часа после обеденного перерыва устраивать регламентированные перерывы продолжительностью 20 минут каждый или продолжительностью 15 минут через каждый час работы.

Продолжительность непрерывной работы на ПК без регламентированного перерыва не должна превышать 2 часа.

Эффективными являются нерегламентированные перерывы (микропаузы) длительностью 1-3 минуты.

Регламентированные перерывы и микропаузы целесообразно использовать для выполнения комплекса упражнений и гимнастики для глаз, пальцев рук, а также массажа. Комплексы упражнений целесообразно менять через 2-3 недели.

Продолжительность рабочего дня не должна быть меньше указанного времени в договоре, но не больше 40 часов в неделю. Для работников до 16 лет – не более 24 часов в неделю, от 16 до 18 лет и инвалидов I и II группы – не более 35 часов.

При работе в ночное время продолжительность рабочей смены сокращается на один час. К работе в ночную смену не допускаются беременные женщины; работники, не достигшие возраста 18 лет; женщины, имеющие детей в возрасте до трех лет, инвалиды, работники, имеющие детей-инвалидов, а также работники, осуществляющие уход за больными членами их семей в соответствии с медицинским заключением, матери и отцы-одиночки детей до пяти лет.

Организация обязана предоставлять ежегодный отпуск продолжительностью 28 календарных дней. Дополнительные отпуска предоставляются работникам, занятым на работах с вредными или опасными условиями труда, работникам имеющими особый характер работы, работникам 64 с ненормированным рабочим днем и работающим в условиях Крайнего Севера и приравненных к нему местностях. [41]

7.4.2. Организационные мероприятия при компоновке рабочей зоны

Рабочее место должно быть организовано в соответствии с требованиями стандартов, технических условий и (или) методических указаний по безопасности труда. Оно должно удовлетворять следующим требованиям:

- обеспечивать возможность удобного выполнения работ;
- учитывать физическую тяжесть работ;
- учитывать размеры рабочей зоны и необходимость передвижения в ней работающего;
- учитывать технологические особенности процесса выполнения работ.

Невыполнение требований к расположению и компоновке рабочего места может привести к получению работником производственной травмы или развития у него профессионального заболевания. Рабочее место программиста должно соответствовать требованиям СанПин 2.2.2/2.4.1340-03.

Конструкция оборудования и рабочего места при выполнении работ в положении сидя должна обеспечивать оптимальное положение работающего, которое достигается регулированием высоты рабочей поверхности, высоты сидения, оборудованием пространства для размещения ног и высотой подставки для ног. Схемы размещения рабочих мест с персональными компьютерами должны учитывать расстояния между рабочими столами с мониторами: расстояние между боковыми поверхностями мониторов не менее 1,2 м, а расстояние между экраном монитора и тыльной частью другого монитора не менее 2,0 м. Клавиатура должна располагаться на поверхности стола на расстоянии 100-300 мм от края, обращенного к пользователю. Быстрое и точное считывание информации обеспечивается при расположении плоскости экрана 65° ниже уровня глаз пользователя, предпочтительно перпендикулярно к нормальной линии взгляда (нормальная линия взгляда 15° градусов вниз от горизонтали). Рабочие места с компьютерами при выполнении творческой работы, требующей значительного умственного напряжения или высокой концентрации внимания, рекомендуется изолировать друг от друга перегородками высотой 1,5 - 2,0 м.

ЗАКЛЮЧЕНИЕ

В результате выполнения выпускной квалификационной работы была реализована модель прогнозирования на основе штрафных сплайнов. Модель разработана на языке C# и представлена в виде десктоп-приложения с использованием технологии WPF. Разработка приложения производилась в среде Microsoft Visual Studio 2018.

Высокое качество прогнозирования обеспечивает метод оптимизации входных параметров сплайна генетическим алгоритмом. Его применение объясняется соотношением качества результатов оптимизации и скоростью использования.

Полученные результаты оценки качества прогнозирования для модели, основанной на штрафных сплайнах, с применением алгоритма оптимизации находятся в пределах допустимых погрешностей и примерно соотносятся с результатами оценок качества для других популярных методов прогнозирования. При этом данная модель обеспечивает возможность прогнозирования в масштабе реального времени.

Разработанное для применения данной модели прогнозирования приложение обеспечивает возможность загрузки пользовательских данных, получение прогнозных данных на их основе, а также получения оптимальных входных параметров. Данное приложение может применяться как в образовательных и исследовательских целях при анализе используемой модели прогнозирования, так и при необходимости получить прогноз для реального набора данных.

CONCLUSION

As a result of the final qualifying work was implemented forecasting model based on penalty splines. The model is developed with C# and implemented as a desktop application using WPF technology. The application was developed in Microsoft Visual Studio 2018.

The high quality of forecasting provides an optimizing method of spline input parameters by genetic algorithm. Its usage is explained by the quality of the optimization results and the speed of use.

The obtained results of forecasting quality evaluation for the model based on penalty splines using the optimization algorithm are within acceptable errors and roughly correlate with the results of quality estimates for other popular forecasting methods. At the same time, this model provides the ability to forecast in real time scale.

The application developed for the usage of this forecasting model provides the ability of loading user data, obtaining forecast data based on them, as well as obtaining optimal input parameters. This application can be used both for educational and research purposes in the analysis of the used forecasting model, and if necessary, to obtain a forecast for a real data set.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Montgomery D. C., Jennings C. L., Kulahci M. Introduction to time series analysis and forecasting. – New Jersey: Wiley, 2015. – 655 p.
2. Zhengtang L. Short-term wind power combined forecasting based on error forecast correction // Energy Conversion and Management. – 2016. – №119. – P. 215-226.
3. Sbrana, G., Silvestrini, A., Venditti, F. Short-term inflation forecasting: The M.E.T.A. approach // Short-term inflation forecasting: The M.E.T.A. approach. – 2017. – №33. – P. 1065-1081.
4. Кондратьев М. А. Методы прогнозирования и модели распространения заболеваний // Компьютерные исследования и моделирование. – 2013. – №5. – С. 863-882.
5. Ермакова Е. А. Методические подходы к прогнозированию тенденций угроз природного и техногенного характера //
6. Тихонов Э. Е. Прогнозирование в условиях рынка. – Невинномысск, 2006. – 221 с.
7. Чучуева И. А. Модель прогнозирования временных рядов по выборке максимального подобия [Электронный ресурс] : дис. ... канд. тех. наук / Московский государственный технический университет им. Н. Э. Баумана. Москва, 2012.
8. Jingfei Yang M. Sc. Power System Short-term Load Forecasting: Thesis for Ph.d degree. Germany, Darmstadt, Elektrotechnik und Informationstechnik der Technischen Universität, 2006. 139 p.
9. Лукашин Ю. П. Адаптивные методы прогнозирования экономических показателей. – М.: Финансы и статистика, 2003. – 416 с.
10. Светуньков И. С., Светуньков С. Г. Методы социально-экономического прогнозирования: учебник и практикум для академического бакалавриата. В 2 т. Том 2. Модели и методы. – М.: Издательство Юрайт, 2017. – 447 с.

11. Box G. P., Jenkins G. M. Time series analysis forecasting and control. – San Francisco: Holden-Day, 1970. – 554 p.
12. Кратович П. В. Нейронные сети и модели Arima для прогнозирования котировок // Программные продукты и системы. 2011. №1. URL: <https://cyberleninka.ru/article/n/neyronnye-seti-i-modeli-arima-dlya-prognozirovaniya-kotirovok> (дата обращения: 04.03.2018).
13. Voyant C., Notton G., Darras C. Uncertainties in global radiation time series forecasting using machine learning: The multilayer perceptron case // Energy – 2017. – №15. – P. 248-257.
14. Srivastana S., Lessmann S. A comparative study of LSTM neural networks in forecasting day-ahead global horizontal irradiance with satellite data // Solar Energy – 2018. – №162. – P. 232-247.
15. Та, X., Wei, Y. Research on a dissolved oxygen prediction method for recirculating aquaculture systems based on a convolution neural network // Computers and Electronics in Agriculture – 2018. – №145. – P. 302-310.
16. Тимошенко П.Н., Яковенко В.С. Экономические циклы – новые подходы к обнаружению, анализу, прогнозированию. – Ставрополь: Изд-во СГАУ, 2003. – 103 с.
17. Кочегурова Е.А., Шебеко Е.В. Использование вариационного сглаживающего сплайна в задаче краткосрочного прогнозирования // Известия ТПУ. 2006. №7. URL: <https://cyberleninka.ru/article/n/ispolzovanie-variatsionnogo-sglazhivayuschego-splayna-v-zadache-kratkosrochnogo-prognozirovaniya> (дата обращения: 01.03.2018).
18. Завьялов Ю.С., Квасов Б.И., Мирошниченко В.Л. Методы сплайн-функций. –М.: Наука, 1980. – С. 352.
19. Низамитдинов А. И. Использование штрафного сплайна (P-spline) и кубического сглаживающего сплайна в прогнозировании временных рядов // Вестник ТГУПБП. 2012. №4 (52). URL: <https://cyberleninka.ru/article/n/ispolzovanie-shtrafnogo-splayna-p-spline-i->

kubicheskogo-sglazhivayuschego-splayna-v-prognozirovanii-vremennyh-ryadov
(дата обращения: 11.03.2018).

20. Кочегурова Е.А., Горохова Е.С. Текущее оценивание производной нестационарного процесса на основе рекуррентного сглаживающего сплайна // Автометрия – 2016. – №3. – С. 79-85.

21. Кочегурова Е.А., Кочегуров А.И., Рожкова Н.Е. Частотный анализ рекуррентных вариационных Р-сплайнов // Автометрия – 2017. – №6. – С. 67-76.

22. Schherbakov M., Brebels A., Schherbakova N., Tyukov A. A Survey of Forecast Error Measures // World Applied Sciences Journal (WASJ). - 2013. - Vol. 24, Spec. Issue 24 : Information Technologies in Modern Industry, Education & Society. – P. 171-176.

23. Schherbakov M., Brebels A. Outliers and anomalies detection based on neural networks forecast procedure: Proceedings of the 31st Annual International Symposium on Forecasting (ISF 2011) // Prague: International Institute of Forecasters, 2011. – P. 21-22.

24. Захарова Е.М., Минашина И.К. Обзор методов многомерной оптимизации // Информационные процессы. – 2014. – №3(14). – С.256-274.

25. Барабашова О.В., Крушель Е.Г. Алгоритмы поиска экстремума функции многих переменных. Методические указания – Волгоград: гос. техн. ун-т, Волгоград, 2000. – 30 с.

26. Спицын В.Г., Цой Ю.Р. Интеллектуальные системы. Учебное пособие – Томск: Изд-во Томского политехнического университета, 2012. – 176 с.

27. Троелсен Эндрю. Язык программирования C# 5.0 и платформа .NET 4.5, 6-е изд. – М. : ООО «И.Д.Вильямс», 2013 – 1312 с.

28. Введение в WPF // MSDN URL: [https://msdn.microsoft.com/ru-ru/library/aa970268\(v=vs.100\).aspx](https://msdn.microsoft.com/ru-ru/library/aa970268(v=vs.100).aspx) (Дата обращения 13.05.2018)

29. Руководство Microsoft по проектированию архитектуры приложений. 2е издание. / URL: https://blogs.technet.microsoft.com/isv_team/2010/09/16/microsoft-patterns-practices/ (дата обращения: 13.05.2018).

30. Какие бывают требования? [Электронный ресурс]: UML2.ru Сообщество аналитиков // <http://www.uml2.ru> URL: <http://www.uml2.ru/faq/faq-requirements/34/> (дата обращения: 13.05.2018)

31. UML — диаграмма вариантов использования (use case diagram) [Электронный ресурс]: Хабрахабр // <https://habrahabr.ru> URL: <https://habrahabr.ru/post/47940/> (дата обращения: 13.05.2018)

32. Охрана труда [Электронный ресурс] / Безопасность жизнедеятельности. URL: <http://www.grandars.ru/shkola/bezopasnost-zhiznedeyatelnosti/ohrana-truda.html>, свободный (дата обращения: 05.05.2018).

33. СанПиН 2.2.2/2.4.1340-03 Гигиенические требования к персональным электронно-вычислительным машинам и организации работы.

34. СанПиН 2.2.4.548-96 Гигиенические требования к микроклимату производственных помещений.

35. ГОСТ 12.1.038–82 Система стандартов безопасности труда. Электробезопасность. Предельно допустимые значения напряжений прикосновения и токов.

36. Федеральный закон от 22.07.2008 N 123-ФЗ (ред. От 13.07.2015) «Технический регламент о требованиях пожарной безопасности» [Электронный ресурс] / КонсультантПлюс. URL: http://www.consultant.ru/document/cons_doc_LAW_78699/, свободный (дата обращения: 05.05.2018).

37. Ефремова О. С. Требования охраны труда при работе на персональных электронно-вычислительных машинах. – 2-е изд., перераб. и доп. – М. : Издательство «Альфа-Пресс», 2008.

38. Назаренко О. Б. Безопасность жизнедеятельности: учебное пособие / О. Б. Назаренко, Ю. А. Амелькович; Томский политехнический университет.

– 3-е изд., перераб. и доп. – Томск: Изд-во Томского политехнического университета, 2013.

39. НПБ 105-03 Определение категорий помещений, зданий и наружных установок по взрывопожарной и пожарной опасности [Электронный ресурс] / Электронный фонд правовой и нормативно-технической документации. URL: <http://docs.cntd.ru/document/1200032102>, свободный (дата обращения: 05.05.2018).

40. ППБ 01–03. Правила пожарной безопасности в Российской Федерации. – М.: Министерство Российской Федерации по делам гражданской обороны, чрезвычайным ситуациям и ликвидации последствий стихийных бедствий, 2003.

41. Трудовой кодекс Российской Федерации" от 30.12.2001 N 197-ФЗ [Электронный ресурс] / Консультант Плюс. URL: https://www.consultant.ru/document/cons_doc_LAW_34683/, свободный (дата обращения: 05.05.2018).

ПРИЛОЖЕНИЕ А

Достоинства и недостатки методов оптимизации

Метод поиска по симплексу	
<p>Достоинства:</p> <ul style="list-style-type: none"> • простота • малое количество заранее установленных параметров • эффективный при ошибках 	<p>Недостатки:</p> <ul style="list-style-type: none"> • трудности при масштабировании задачи • низкая скорость выполнения • невозможно изменить размер симплекса без пересчета всех значений
Метод Хука-Дживса	
<p>Достоинства:</p> <ul style="list-style-type: none"> • простая стратегия поиска • малый объем требуемой памяти 	<p>Недостатки:</p> <ul style="list-style-type: none"> • возможность вырождения алгоритма • заикливание
Метод Ньютона (алгоритм наискорейшего спуска с шагом, длина которого зависит от свойств минимизируемой функции)	
<p>Достоинства:</p> <ul style="list-style-type: none"> • поиск решений в некоторых случаях можно свести к одному шагу (если минимизируемая функция квадратичная или относится к классу поверхностей вращения) 	<p>Недостатки:</p> <ul style="list-style-type: none"> • необходимость вычисления вторых производных • возможность появления значительных вычислительных погрешностей
Нейронная сеть Хопфилда	
<p>Достоинства:</p> <ul style="list-style-type: none"> • Возможность применения в нестандартных задачах • Высокая скорость выполнения при предварительном обучении 	<p>Недостатки:</p> <ul style="list-style-type: none"> • Высокая вероятность застревания в локальном экстремуме
Генетический алгоритм	
<p>Достоинства</p> <ul style="list-style-type: none"> • Применяется для решения сложных неформальных задач • Обеспечение решения нестандартных проблем 	<p>Недостатки</p> <ul style="list-style-type: none"> • Отсутствие гарантии того, что будет обнаружено глобальное решение за приемлемое время
Интервальный метод ветвей и границ	
<p>Достоинства</p> <ul style="list-style-type: none"> • Возможность определения глобального экстремума 	<p>Недостатки:</p> <ul style="list-style-type: none"> • Большие затраты времени при решении задач большой размерности

ПРИЛОЖЕНИЕ Б

Описание классов программного приложения

Наименование	Краткое описание
InputData	Хранит информацию о входных данных пользователя, а именно о входных параметрах сплайна и данные, предназначенные для дальнейшего прогнозирования
InputFunctionData	Хранит информацию о входных данных тестовой функции и параметрах сплайна
ExcelData	Модель представления данных при экспорте в Excel
Forecasting	Хранит в себе алгоритмы прогнозирования, общие для дочерних классов DataForecasting и FunctionForecasting
DataForecasting	Вычисление прогнозных значений для пользовательских данных
FunctionForecasting	Вычисление значений сплайна для тестовых функций
Errors	Хранит в себе поля и методы, требуемые для вычисления различных видов погрешностей для оценки качества прогнозирования
Parsing	Реализует парсинг файла с входными пользовательскими данными
Person	Представляет особь для реализации ГА и хранит в себе данные о хромосомном наборе конкретной особи и значение приспособленности
Population	Представляет собой поколение особей для реализации ГА и содержит массив особей одного поколения
EvaluatingPerson	Реализует оценивание особей с использованием функции приспособленности
GeneticAlgorithm	Хранит в себе реализацию генетического алгоритма для решения задачи оптимизации
ForecastingWindow	Пользовательский интерфейс

ПРИЛОЖЕНИЕ В

Матрица SWOT-анализа

	<p>Сильные стороны научного исследования:</p> <p>С1. Понятный и «дружественный» пользовательский интерфейс</p> <p>С2. Простота эксплуатации</p> <p>С3. Небольшие затраты ресурсов ОС</p>	<p>Слабые стороны научного исследования:</p> <p>Сл1. Ограниченность функциональных возможностей системы</p> <p>Сл2. Невозможность пользовательской настройки интерфейса</p> <p>Сл3. Неизвестность используемой модели прогнозирования на целевом рынке</p>
<p>Возможности научного исследования:</p> <p>В1. Заинтересованность в разработке смежных сегментов рынка</p> <p>В2. Использование разработки в образовательных целях</p> <p>В3. Увеличение популярности прогнозирования в различных сферах производства.</p>	<p>Направления развития:</p> <p>1. В1В2В3С1С2 – модернизация приложения, визуализация результатов, отсутствие «минимального порога» вхождения</p> <p>2. В2С3 – оптимизация приложения для уменьшения затрат потребляемой памяти и ресурсов операционной системы</p>	<p>Сдерживающие факторы:</p> <p>1. В1Сл1Сл2 – невозможность кастомизации приложения под нужды конкретного пользователя</p> <p>2. В2Сл3 – отсутствие популярных научных исследований используемой модели прогнозирования</p> <p>3. В3Сл2Сл3 – отсутствие популярности разработки на рынке</p>
<p>Угрозы научного исследования:</p> <p>У1. Появление систем-конкурентов</p> <p>У2. Появление новых моделей прогнозирования с более высоким уровнем точности</p> <p>У3. Потеря интереса клиентов к десктоп-приложениям</p>	<p>Угрозы развития:</p> <p>1. В1Сл1Сл2 – отсутствие интереса представителей смежных областей к данному приложению</p> <p>2. В3Сл2Сл3 – Возникновение интереса к другим видам прогнозирования и другим типам приложений</p>	<p>Уязвимости:</p> <p>1. У1Сл1Сл2Сл3 – отсутствие перспективных преимуществ перед компаниями-конкурентами</p>

ПРИЛОЖЕНИЕ Г

Перечень этапов, работ и распределение исполнителей

Основные этапы	№ раб	Содержание работ	Должность исполнителя
1	2	3	4
Разработка технического задания	1	Составление и утверждение технического задания	Научный руководитель
Выбор направления исследований	2	Принятие решения о разработке, постановка целей и задач, получение исходных данных	Научный руководитель
	3	Подбор и изучение материалов по тематике	Студент
	4	Календарное планирование работ по теме	Научный руководитель, студент
Теоретические и экспериментальные исследования	5	Изучение существующих технологий краткосрочного прогнозирования	Студент
	6	Изучение используемой в разработке модели прогнозирования	Студент
	7	Определение структуры разрабатываемого ПО, выбор средств разработки	Студент
	8	Работа над отдельными модулями программы	Студент
	9	Редактирование и отладка программы	Студент
Обобщение и оценка результатов	10	Анализ результатов работы	Научный руководитель, студент
	11	Выбор метода оптимизации входных параметров	Научный руководитель, студент
	12	Анализ результатов после внедрения оптимизации	Научный руководитель, студент
Разработка технической документации и проектирование	13	Выпуск документации по программному приложению	Студент
	14	Оформление графического материала, визуализация данных	Студент
	15	Оценка эффективности и применения разработки	Студент
	16	Разработка социальной ответственности по теме	Студент

Оформление отчета	17	Составление пояснительной записки (эксплуатационно-технической документации)	Студент
-------------------	----	--	---------

ПРИЛОЖЕНИЕ Д

Временные показатели научного исследования

Название работы	Исполнители	Трудоемкость работ, человеко-дни									Длительность работ					
		tmin			tmax			toji			Тр, рабочие дни			Тк, календарные дни		
		И1	И2	И3	И1	И2	И3	И1	И2	И3	И1	И2	И3	И1	И2	И3
Составление и утверждение технического задания	1	1	1	1	2	2	2	1,4	1,4	1,4	1,4	1,4	1,4	2	2	2
Принятие решения о разработке, постановка целей и задач, получение исходных данных	1	1	1	1	3	3	3	1,8	1,8	1,8	1,8	1,8	1,8	3	3	3
Подбор и изучение материалов по тематике	1	5	6	6	10	11	11	7	8	8	7	8	8	10	12	12
Календарное планирование работ по теме	2	1	1	1	2	2	2	1,4	1,4	1,4	0,7	0,7	0,7	1	1	1
Изучение существующих технологий краткосрочного прогнозирования	1	10	10	10	15	15	15	12	12	12	12	12	12	18	18	18
Изучение используемой в разработке модели прогнозирования	1	7	7	7	10	10	10	8,4	8,4	8,4	8,4	8,4	8,4	12	12	12
Определение структуры разрабатываемого ПО,	1	2	5	3	3	6	4	2,4	5,4	3,4	2,4	5,4	3,4	4	8	5

выбор средств разработки																	
Работа над отдельными модулями программы	1	15	16	21	18	20	25	16,2	17,6	22,6	16,2	17,6	22,6	24	26	33	
Редактирование и отладка программы	1	4	8	5	6	10	7	4,8	8,8	5,8	4,8	8,8	5,8	7	13	9	
Анализ результатов работы	2	4	4	4	6	6	6	4,8	4,8	4,8	2,4	2,4	2,4	4	4	4	
Выбор метода оптимизации входных параметров, анализ результатов после внедрения оптимизации	2	3	3	3	4	4	4	3,4	3,4	3,4	1,7	1,7	1,7	3	3	3	
Выпуск документации по программному приложению	1	2	2	2	4	4	4	2,8	2,8	2,8	2,8	2,8	2,8	4	4	4	
Оформление графического материала, визуализация данных	1	2	2	2	3	3	3	2,4	2,4	2,4	2,4	2,4	2,4	4	4	4	
Оценка эффективности и применения разработки	1	3	3	3	5	5	5	3,8	3,8	3,8	3,8	3,8	3,8	6	6	6	
Разработка социальной ответственности по теме	1	3	3	3	5	5	5	3,8	3,8	3,8	3,8	3,8	3,8	6	6	6	

Составление пояснительной записки (эксплуатационно- технической документации)	1	3	3	3	5	5	5	3,8	3,8	3,8	3,8	3,8	3,8	3,8	6	6	6
Итого								80,2	89,6	89,6	75,4	84,8	84,8	114	128	128	