



МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ
федеральное государственное бюджетное образовательное учреждение высшего
образования
«РОССИЙСКИЙ ГОСУДАРСТВЕННЫЙ ГИДРОМЕТЕОРОЛОГИЧЕСКИЙ
УНИВЕРСИТЕТ»

Кафедра Информационные технологии и системы безопасности

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА

(Бакалавр)

На тему «Идентификация состояния устройств бортовых систем на основе
бэггинга»

Исполнитель Николаев Никита Александрович
(фамилия, имя, отчество)

Руководитель доктор технических наук
(ученая степень, ученое звание)

Сухопаров Михаил Евгеньевич
(фамилия, имя, отчество)

**«К защите допускаю»
Заведующий кафедрой**

(подпись)

доктор технических наук
(ученая степень, ученое звание)

Бурлов Вячеслав Георгиевич
(фамилия, имя, отчество)

« » 2023 г.

Санкт-Петербург

2023

Реферат

Выпускная квалификационная работа объемом 62 страницы содержит 13 рисунков, 1 таблицу, 21 источник литературы и 1 приложение.

Объект исследования - устройства бортовых автоматизированных систем.

Предмет исследования – является методы идентификации состояния устройств и узлов бортовых автоматизированных систем.

Целью работы является разработка метода идентификации состояния устройств бортовых автоматизированных систем при помощи методов машинного обучения.

Задачи:

1. Анализ теоретической базы методов машинного обучения и их применения в задачах идентификации состояния.

2. Изучение специфики работы устройств бортовых автоматизированных систем и анализ собранных данных.

3. Разработка метода идентификации состояния устройств и узлов бортовых автоматизированных систем на основе методов машинного обучения.

4. Проведение анализа данных с использованием выбранного метода и оценка его применимости.

5. Анализ полученных результатов

В выпускной квалификационной работе произведено исследование существующих методов и подходов к идентификации состояния устройств бортовых систем. В дальнейшем метод бэггинга был использован для идентификации устройств бортовых систем.

Содержание

СПИСОК ТЕРМИНОВ И СОКРАЩЕНИЙ	4
ВВЕДЕНИЕ.....	5
ГЛАВА 1. ИДЕНТИФИКАЦИЯ СОСТОЯНИЯ БОРТОВЫХ СИСТЕМ	8
1.1 Классификация бортовых систем.....	8
1.2 Оценка состояния процессов функционирования бортовых систем.....	15
1.3 Постановка задачи.....	20
Вывод.....	22
ГЛАВА 2. ПРИМЕНЕНИЕ БЭГГИНГА В МОДЕЛЯХ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ИДЕНТИФИКАЦИИ СОСТОЯНИЯ БОРТОВЫХ СИСТЕМ	24
2.1 Методы машинного обучения и способы повышения качества обработки информации для оценки состояния бортовых систем.....	24
2.2 Применение бэггинга в методах МО для повышения качества оценки состояния бортовых систем	38
ГЛАВА 3. АНАЛИЗ ПРИМЕНИМОСТИ ПРЕДЛАГАЕМОГО ПОДХОДА	46
3.1 Описание предлагаемого подхода.....	46
3.2 Оценка применимости предлагаемого подхода.....	49
ЗАКЛЮЧЕНИЕ	57
Источники	59

СПИСОК ТЕРМИНОВ И СОКРАЩЕНИЙ

МО - машинное обучение

МИС - морские информационные системы

ИНС - искусственные нейронные сети

ВВЕДЕНИЕ

Актуальность выпускной квалификационной работы. В современном мире технологии играют важную роль в различных сферах жизни, включая транспортную систему и, конечно, судоходство. Современные суда оснащены множеством различных устройств и систем, которые помогают обеспечить безопасность и комфорт пассажиров, членов экипажа, сохранность грузов, и, в целом, осуществлять контроль движения судов в соответствии с международными правилами и нормами. Развитие судостроения вызывает необходимость оценки работоспособности, функциональной безопасности отдельных бортовых устройств. Однако, недостаточное внимание к состоянию этих устройств может привести к авариям и несчастным случаям. Повсеместная автоматизация технологических процессов не оставила в стороне и морскую отрасль – современные суда оснащены множеством вспомогательных средств контроля мореходности, связи, навигации и т.д., которые, зачастую, объединены в единую бортовую автоматизированную информационную систему. С течением времени подобные системы становятся все сложнее, растет количество обрабатываемых данных.

Таким образом, существует объективное противоречие между возможностями, которые предоставляют современные бортовые автоматизированные системы, и существующим научно-методическим и математическим обеспечением этих систем, реализующих алгоритмы автоматизированной обработки данных в целях выявления различных инцидентов. Следствием неразрешенности этого противоречия является объективная необходимость развития и адаптации методов математического обеспечения специализированных информационных систем (ИС), интегрируемых в бортовые автоматизированные комплексы, в целях противодействия внешним и внутренним деструктивным воздействиям, а также сбоям в процессе функционирования. Таким образом, обоснование и

разработка теоретических основ математического и программного обеспечения систем мониторинга информации, направленных на разрешение противоречий между возрастающими потребностями повышения показателей устойчивости, безопасности функционирования объектов морской техники и традиционными способами анализа состояния бортовых систем, является **актуальной научной проблемой**.

В ходе функционирования устройств и узлов бортовых автоматизированных систем могут возникать сбои как на уровне информационной системы, так и отдельного устройства. В целях предотвращения подобных инцидентов происходит совершенствование и адаптация моделей, методов мониторинга состояния, направленных на оценку функциональности, безопасности и производительности. В их основе лежат принципы статистического анализа, анализа причинно-следственных связей, переходов, формирования прецедентных, событийных моделей. Модели, основанные на статистике, накапливают информацию о параметрах функционирования в различных режимах и состояниях, что в дальнейшем позволяет применять различные методы машинного обучения с целью выявления аномальных ситуаций, возникающих в процессе работы объектов морской техники.

Объектом исследования являются устройства бортовых автоматизированных систем.

Предметом исследования является методы идентификации состояния устройств и узлов бортовых автоматизированных систем.

Целью данной работы является разработка метода идентификации состояния устройств бортовых автоматизированных систем при помощи методов машинного обучения.

Для достижения указанной цели решались следующие **научно-технические задачи**:

1. Анализ теоретической базы методов машинного обучения и их применения в задачах идентификации состояния.

2. Изучение специфики работы устройств бортовых автоматизированных систем и анализ собранных данных.

3. Разработка метода идентификации состояния устройств и узлов бортовых автоматизированных систем на основе методов машинного обучения.

4. Проведение анализа данных с использованием выбранного метода и оценка его применимости.

5. Анализ полученных результатов.

Для решения поставленных задач в выпускной квалификационной работе использовались теории вероятностей и математической статистики, классификации, защиты информации, надежности, методы машинного обучения и нейросетевые технологии. Проводился анализ государственных стандартов в области надежности и информационной безопасности, была произведена классификация современных бортовых автоматизированных систем. Разработан метод идентификации состояния устройств бортовых систем, произведена оценка его применимости, сделаны выводы в рамках полученных результатов.

ГЛАВА 1. ИДЕНТИФИКАЦИЯ СОСТОЯНИЯ БОРТОВЫХ СИСТЕМ

В первой главе выпускной квалификационной работы проводится анализ тенденций развития современных бортовых автоматизированных систем, проводится их классификация. Проводится оценка состояния процессов функционирования бортовых систем, рассматриваются государственные стандарты, связанные с надежностью и информационной безопасностью. Проводится общая постановка задачи исследования.

1.1 Классификация бортовых систем

Бортовая система — это комплекс оборудования и программного обеспечения, установленный на борту транспортного средства, который обеспечивает контроль и управление его работой. Бортовая система может включать в себя системы навигации, автоматического управления, контроля двигателя, системы безопасности и коммуникационные системы. Она позволяет повышать эффективность и безопасность работы транспортного средства, а также улучшать комфорт и удобство его использования.

Как именно оснащено судно, зависит от его назначения и типа. Некоторые из возможных систем, которыми может быть оснащено судно, включают в себя: системы навигации, системы связи, системы безопасности, системы управления двигателями, системы питания и электроснабжения, системы водоочистки и канализации, системы охлаждения и кондиционирования воздуха, системы пожаротушения, системы спасательных средств и другие.

Какими бортовыми системами оснащено судно зависит от типа и назначения судна. Однако, в общей сложности на судне могут быть установлены следующие системы [9]:

1. Система управления движением и маневренности судна.

Система управления движением и маневренности судна включает в себя различные компоненты и устройства, которые позволяют капитану и экипажу эффективно управлять судном в различных условиях и ситуациях.

Основными элементами системы управления движением и маневренности судна являются:

- Рулевое устройство - устройство, которое позволяет изменять направление движения судна. Рулевое устройство может быть механическим или гидравлическим.

- Двигательное устройство - устройство, которое обеспечивает движение судна. Двигательное устройство может быть дизельным, газотурбинным или электрическим.

- Пропульсивное устройство - устройство, которое преобразует энергию двигателя в движение судна. Пропульсивное устройство может быть винтом, водометом или гидродинамическим.

- Система управления движением - система, которая обеспечивает контроль и регулирование скорости и направления движения судна. Система управления движением может быть автоматической или ручной.

- Система маневренности - система, которая обеспечивает возможность судна выполнять различные маневры, такие как повороты, остановки и т.д. Система маневренности может включать в себя буксирные устройства, якорные устройства и т.д.

2. Системы безопасности.

Системы безопасности на судне включают в себя следующие элементы:

- Системы пожаротушения. Включают в себя пожарные извещатели, огнетушители, пожарную автоматику, системы пены и газового пожаротушения.

- Системы спасательного оборудования. Включают в себя спасательные шлюпки, спасательные жилеты, спасательные круги, спасательные плоты, а также системы аварийного сигнализирования и связи.

- Системы безопасности персонала. Включают в себя системы охраны труда, медицинские центры, системы обучения и тренировки персонала в случае аварийных ситуаций.

3. Системы связи и навигации (радиооборудование, радары, GPS, компьютерные системы управления и навигации).

Системы связи и навигации на судне включают в себя множество устройств и оборудования, которые обеспечивают связь между судном и внешним миром, а также позволяют определить местоположение и направление движения судна.

Основными системами связи на судне являются радиосвязь, спутниковая связь, радиолокационная связь и системы сигнализации. Радиосвязь позволяет общаться с другими судами, береговыми станциями и спасательными службами. Спутниковая связь обеспечивает связь в отдаленных районах и на больших расстояниях от берега. Радиолокационная связь используется для обнаружения других судов и определения их расстояния и направления.

Системы навигации на судне включают в себя GPS-навигацию, эхолот, гироскопы, компасы и другие устройства. GPS-навигация позволяет определять местоположение судна в режиме реального времени. Эхолот используется для определения глубины и контуров морского дна. Гироскопы и компасы используются для определения направления движения судна и его ориентации в пространстве.

Важность систем связи и навигации на судне заключается в том, что они обеспечивают безопасность плавания и защиту жизни и здоровья экипажа и пассажиров. Они также позволяют экипажу судна эффективно управлять им и выполнить задачи, связанные с его движением и грузоперевозками.

4. Системы питания и энергоснабжения.

Системы питания и энергоснабжения на судне состоят из нескольких компонентов:

- Генераторы: основной и запасной. Они производят электрическую энергию для всего судна.
- Аккумуляторные батареи: они хранят электрическую энергию, которая может быть использована во время отключения генераторов.
- Системы распределения электроэнергии: они обеспечивают электропитание для всех систем и оборудования на судне.
- Системы автоматического управления и контроля: они следят за работой генераторов и батарей, и автоматически включают запасные системы в случае отказа основных.
- Системы электрогенерации: они позволяют использовать движение судна для производства электрической энергии, например, с помощью гидравлических турбин.
- Системы кондиционирования и обогрева: они обеспечивают комфортные условия для экипажа и пассажиров.

Все эти системы работают вместе, чтобы обеспечить надежное энергоснабжение на судне.

5. Системы обработки воды и канализации.

Система обработки воды и канализации на судне имеет основное назначение – очистка и переработка отходов воды и канализации, чтобы снизить негативное воздействие на окружающую среду. Существуют различные системы обработки воды и канализации на судне, но в основном они включают следующие компоненты:

- Система очистки сточных вод. Эта система включает в себя физическую, химическую и биологическую очистку сточных вод, чтобы устранить загрязнения, бактерии и вирусы, которые могут быть опасными для здоровья людей и окружающей среды.
- Система очистки и переработки питьевой воды. Эта система обеспечивает постоянный доступ к чистой и безопасной питьевой воде на судне. Она включает в себя фильтрацию, дезинфекцию, умягчение и дистилляцию воды.

- Система сброса отходов. Эта система позволяет правильно утилизировать отходы на судне, включая пищевые отходы, бумажные отходы, стекло, пластик и металл.

- Система сброса канализации. Эта система обеспечивает безопасный сброс обработанной канализации в море. Она включает в себя фильтрацию, обеззараживание и сброс воды.

Все системы обработки воды и канализации на судне должны соответствовать международным нормативным требованиям и стандартам, чтобы минимизировать воздействие на окружающую среду и защитить здоровье людей, работающих на судне.

7. Системы управления грузоподъемностью.

На судне используются различные системы управления грузоподъемностью, которые обеспечивают безопасную и эффективную работу с грузами.

Одной из ключевых систем является грузовой кран. Он может быть установлен как на палубе, так и на борту судна. Грузовые краны могут быть различных типов, включая тросовые, гидравлические и пневматические. Краны обычно оборудованы системой автоматического контроля нагрузки, которая обеспечивает безопасную работу крана и предотвращает перегрузку.

Другой важной системой является система баллаستирования. Она используется для управления распределением воды в балластных танках, что позволяет контролировать стабильность и грузоподъемность судна. Система балластирования позволяет поддерживать оптимальный уровень стабильности судна в зависимости от условий на море и грузового веса.

Также на судне может быть установлена система динамической позиционирования. Она обеспечивает автоматическое управление двигателями и рулевым управлением, чтобы удерживать судно на заданной позиции без использования якоря. Это позволяет судну стабильно работать в условиях сильного течения или ветра.

Наконец, система обнаружения перегрузки также может быть установлена на судне. Она позволяет мгновенно определить перегрузку груза и предупредить экипаж о необходимости срочных мер. Это обеспечивает безопасность грузовой работы и предотвращает повреждение судна.

8. Системы контроля и диагностики.

На судне используются различные системы контроля и диагностики для обеспечения безопасности и эффективности его работы. Некоторые из них:

- Система управления и контроля двигателя. Эта система контролирует работу двигателя и его параметры, такие как температура, давление масла, скорость вращения, топливный расход и др.

- Система контроля и диагностики электрооборудования. Она позволяет контролировать состояние и работу электрооборудования на судне, включая генераторы, аккумуляторы, электродвигатели и др.

- Система контроля и диагностики навигационного оборудования. Она позволяет контролировать работу радаров, эхолотов, GPS-навигаторов и других навигационных приборов.

- Система контроля и диагностики системы пожаротушения. Она позволяет контролировать работу системы пожаротушения на судне и обнаруживать возможные неисправности.

- Система контроля и диагностики системы жизнеобеспечения. Она позволяет контролировать работу системы питания, вентиляции, кондиционирования воздуха и других систем, обеспечивающих комфорт и безопасность на судне.

- Система контроля и диагностики системы очистки и утилизации отходов. Она позволяет контролировать работу системы очистки и утилизации отходов на судне и обнаруживать возможные неисправности.

Все эти системы контролируются и диагностируются с помощью специальных приборов и программного обеспечения, что позволяет

своевременно обнаруживать неисправности и предотвращать аварии и происшествия на судне.

9. Системы, связанные с экологической безопасностью (системы очистки отходов, разделения отходов, системы очистки выбросов).

- Система управления отходами (Waste Management System) — это система, которая регулирует управление отходами на судне. Отходы на судне могут быть различными - от продуктов питания до опасных химических веществ. Неконтролируемое складирование отходов на судне может привести к загрязнению морских вод и нарушению экологического баланса. Система управления отходами определяет правила сбора, обработки и утилизации отходов на судне.

- Система контроля выбросов (Emission Control System) - это система, которая контролирует выбросы вредных веществ в атмосферу. Суда могут выбрасывать в атмосферу различные вредные вещества, такие как оксиды азота, серы или углерода. Эти вещества могут привести к загрязнению воздуха и здоровью людей. Система контроля выбросов устанавливает правила контроля выбросов вредных веществ на судне.

Все эти бортовые системы и их устройства взаимосвязаны между собой и у всех этих элементов судна одна цель - сделать судно безопаснее и комфортнее. Кроме того, с помощью этих устройств увеличивается поддержка окружающей среды за счет рационального выброса отходов в воду и контроля выбросов вредных веществ в атмосферу планеты.

К сожалению, все эти устройства хоть и проектируются и изготавливаются с особым контролем к мелочам, но невозможно их сделать вечными и за каждым элементом этой огромной сети бортовых систем нужно осуществлять контроль. Без контроля за устройствами, у судна могут возникнуть проблемы, конечно, может возникнуть незначительная проблема, с которой судно может спокойно эксплуатироваться, а экипаж и пассажиры даже не заметят изменений, но может случиться и обратное - когда судно из-за отсутствия идентификации состояния устройств бортовых систем может

судно становится непригодным для дальнейшей эксплуатации. Чтобы избежать этих последствий на каждом судне установлена бортовая система, осуществляющая контроль всех узлов.

1.2 Оценка состояния процессов функционирования бортовых систем

Оценка состояния процессов функционирования бортовых систем судна является важным аспектом обеспечения безопасности плавания. Для этого используются различные методы и технологии, которые позволяют определить состояние систем и обнаружить возможные неисправности [10].

Одним из методов оценки состояния систем является мониторинг и диагностика, которые позволяют следить за работой систем в режиме реального времени и обнаруживать отклонения от нормы. Для этого могут применяться различные сенсоры, контроллеры и программное обеспечение, которые обеспечивают сбор данных и их анализ.

Другим методом оценки состояния систем является техническое обслуживание и ремонт, которые проводятся с целью предотвращения возможных неисправностей и обеспечения надежной работы систем. Для этого используются различные методы диагностики и испытания, которые позволяют выявить неисправности и подобрать оптимальные методы и средства их устранения.

Также важным аспектом оценки состояния систем является обучение и подготовка персонала, который должен обеспечивать их работу. Для этого проводятся различные тренинги и курсы, которые позволяют освоить необходимые знания и навыки, связанные с обслуживанием и ремонтом систем.

Таким образом, оценка состояния процессов функционирования бортовых систем судна является важным аспектом обеспечения безопасности плавания и требует комплексного подхода, который включает в себя

мониторинг и диагностику, техническое обслуживание и ремонт, а также обучение и подготовку персонала.

Модели надежности - это математические модели, которые описывают вероятность отказа системы или устройства в зависимости от времени эксплуатации. Эти модели используются для оценки надежности и прогнозирования срока службы техники.

Понятие надежности - это свойство системы или устройства, которое определяет его способность сохранять работоспособность в течение определенного времени при заданных условиях эксплуатации. Оценка надежности морской техники является важным аспектом ее проектирования и эксплуатации, так как от нее зависит безопасность на море и экономическая эффективность.

Рассчитать надежность морской техники можно с помощью анализа ее технических характеристик, исторических данных о ее эксплуатации и методов математического моделирования. Вот несколько способов:

1. Анализ технических параметров. Оцените надежность морской техники, изучив ее технические характеристики, такие как мощность, скорость, грузоподъемность, длина и ширина. Оцените, насколько эти параметры соответствуют условиям эксплуатации, и сравните их с параметрами аналогичной техники.

2. Исторический анализ. Изучите исторические данные о эксплуатации морской техники, такие как ее рабочий наработка, степень износа, частота ремонта и поломок. Оцените, как часто возникают проблемы с определенными компонентами, и какие меры были приняты для их устранения.

3. Математическое моделирование. Используйте математические модели для расчета надежности морской техники. Эти модели могут включать в себя статистические данные, такие как распределение вероятности времени наработки на отказ, или физические модели, которые

учитывают влияние различных факторов, таких как вибрации, температура и влажность.

4. Экспертный анализ. Проведите экспертную оценку надежности морской техники, пригласив специалистов в отрасли. Это может включать в себя оценку качества производства, квалификацию операторов, уровень обслуживания и т.д.

Надежность бортовых систем на судне - это способность системы выполнять свои функции без сбоев и отказов на протяжении всего срока эксплуатации судна. В зависимости от функционального назначения и характера работы, бортовые системы могут быть подвержены различным видам состояний:

1. Нормальное состояние - когда система работает без сбоев и выполняет свои функции в соответствии с требованиями.

2. Критическое состояние - когда происходит сбой системы, который может привести к серьезным последствиям для безопасности судна и его экипажа.

3. Аварийное состояние - когда происходит полный отказ системы, что может привести к аварии на судне.

4. Предаварийное состояние - когда происходит сбой системы, который может привести к аварии, но еще не достиг критического уровня.

Для обеспечения надежности бортовых систем на судне необходимо проводить регулярное техническое обслуживание и проверку всех систем на соответствие требованиям безопасности и надежности. В случае обнаружения неисправностей, необходимо незамедлительно принимать меры по их устранению, чтобы избежать возможных аварийных ситуаций.

Оценка состояния процессов функционирования бортовых систем на судне является важной задачей для обеспечения безопасности плавания и сохранности жизни и здоровья людей на борту. Для этого используются различные методы и технические средства, такие как:

1. Системы мониторинга и диагностики (СМиД). Это комплексное оборудование, позволяющее непрерывно контролировать работу всех систем и устройств на судне, а также выявлять и диагностировать возможные неисправности.

2. Визуальный осмотр и проверка состояния оборудования. Этот метод включает в себя регулярный осмотр и проверку состояния всех систем, узлов и устройств на судне, а также их техническое обслуживание и ремонт при необходимости.

3. Испытания и проверки систем и устройств на судне. Для этого могут применяться различные методы, такие как испытание на герметичность, проверка на прочность, испытание на нагрузку и т.д.

4. Анализ данных и статистических показателей. Путем анализа данных о состоянии оборудования и эксплуатации судна можно выявить возможные проблемы и улучшить процессы функционирования бортовых систем.

В целом, оценка состояния процессов функционирования бортовых систем на судне должна проводиться регулярно и комплексно, с использованием различных методов и средств контроля и диагностики. Это позволит обеспечить безопасность плавания и предотвратить возможные аварии и неисправности на борту судна.

Определение устойчивости работоспособности бортовых систем судна является комплексным процессом, который включает в себя следующие этапы:

1. Анализ технической документации на судно и его системы. Необходимо изучить характеристики систем, их назначение, принципы работы, технические характеристики и требования к их эксплуатации.

2. Инспекция бортовых систем судна. Необходимо проверить работоспособность каждой системы и ее компонентов, состояние оборудования, наличие повреждений, износа и других дефектов.

3. Оценка рисков. Необходимо оценить вероятность возникновения неполадок в работе систем и их последствий для судна и экипажа.

4. Разработка мероприятий по обеспечению устойчивости работоспособности систем. На основе результатов анализа и инспекции необходимо разработать мероприятия по устранению выявленных дефектов и профилактике возможных неполадок.

5. Проведение испытаний. После проведения мероприятий необходимо провести испытания систем, чтобы убедиться в их работоспособности и устойчивости.

6. Мониторинг и контроль. Для поддержания устойчивости работоспособности систем необходимо проводить регулярный мониторинг и контроль и проводить необходимые ремонтные работы в случае выявления дефектов.

Все эти этапы должны проводиться в соответствии с международными стандартами и требованиями безопасности судоходства.

Самым важным стандартом в Российской Федерации является ГОСТ. ГОСТ является основным документом, регулирующим качество продукции, услуг и процессов в Российской Федерации. Он устанавливает требования к продукции, услугам, процессам и системам управления, которые должны быть соблюдены производителями и поставщиками, чтобы обеспечить высокое качество и безопасность продукции и услуг.

Для оценки состояния процессов функционирования бортовых систем судна ГОСТы могут быть важными в следующих аспектах:

- ГОСТы, регулирующие качество и безопасность судовых систем и оборудования, могут использоваться для оценки соответствия устройств бортовых систем судна требованиям стандартов. Это может быть полезно при проведении идентификации состояния устройств и определении возможных проблем.

ГОСТ Р 56243-2014 [11] “Внутренний водный транспорт. Система управления безопасностью судов. Требования по обеспечению надежности механизмов”

ГОСТ 32455-2013 [12] “Глобальная навигационная спутниковая система. Морская навигационная аппаратура потребителей. Приемные устройства. Общие требования, методы и требуемые результаты испытаний”

- ГОСТы, которые регулируют процессы управления и контроля качества, могут быть применены для разработки стратегии идентификации состояния устройств бортовых систем судна на основе беггинга. Эти стандарты могут помочь определить процедуры и требования для контроля качества и обеспечения соответствия требованиям.

ГОСТ Р ИСО 10006–2019 [13] “Менеджмент качества. Руководящие указания по менеджменту качества в проектах”

- ГОСТы, которые регулируют оборудование и технические требования, могут использоваться для оценки технической готовности устройств бортовых систем судна к выполнению своих функций. Это может быть особенно полезно при определении возможных проблем и уязвимых мест системы.

В целом, ГОСТы играют важную роль в идентификации состояния устройств бортовых систем судна на основе беггинга, обеспечивая соответствие требованиям стандартов и помогая определить возможные проблемы и уязвимости системы.

1.3 Постановка задачи

В современных условиях управление техническими объектами невозможно без использования различных систем мониторинга и контроля. Задача идентификации состояния устройств бортовых систем судна является актуальной и важной в сфере морского транспорта, поскольку позволяет обеспечить безопасность плавания и предотвращать возможные аварии.

Поставленная цель исследования заключается в разработке метода идентификации состояния устройств бортовых систем судна с использованием методов машинного обучения (МО). Машинное обучение — класс методов искусственного интеллекта, характерной чертой которых

является не прямое решение задачи, а обучение за счёт применения решений множества сходных задач. Для построения таких методов используются средства математической статистики, численных методов, математического анализа, методов оптимизации, теории вероятностей, теории графов, различные техники работы с данными в цифровой форме. Все задачи, решаемые с помощью МО, относятся к одной из следующих категорий:

- задача регрессии – прогноз на основе выборки объектов с различными признаками;
- задача классификации – получение категориального ответа на основе набора признаков;
- задача кластеризации – распределение данных на группы;
- задача уменьшения размерности – сведение большого числа признаков к меньшему для удобства их последующей визуализации.

В рамках выпускной квалификационной работы для достижения поставленной цели исследования решается задача классификации состояния устройств и узлов бортовых автоматизированных систем. Исследование качественных показателей различных классифицирующих алгоритмов зависит от параметров обучающей выборки и структуры, и содержащейся в ней информации. Обычно вопросы формирования обучающих выборок связаны с их сбалансированностью, разделением и правильной интерпретацией фоновых и значимых закономерностей, отсутствием обучающих объектов определенного вида, элементов признаковой системы, не точными диапазонами значений переменных, разбалансировкой, появлением внешних закономерностей, обусловленных особенностями формирования обучающего множества. Однако в случае наличия в ней информации разнородных событий возникает ряд проблемных моментов интерпретации структуры информации. В задачах классификации объекты наблюдения связаны с набором меток $\{l_1, l_2, \dots, l_n\} \subseteq L$. Присутствуют события, которые относятся к безопасным и опасным состояниям.

Алгоритмы классификации идентифицируют события, записи о которых помечены меткой l_i из набора непересекающихся меток L , $|L| > 1$. Если $|L| = 2$, то осуществляется бинарная классификация, а если $|L| > 2$ – многоклассовая классификация.

Таким образом, при построении классификаторов применяются методы, преобразующие задачу классификации с несколькими метками в одну или несколько задач классификации с одной меткой, и методы, которые расширяют возможности для обработки данных с несколькими метками.

Рассматривается конечное число меток l_i множества L . Каждая метка l_i входит в подмножества L_1 или L_2 множества L , определяющие опасные C_1 и безопасные C_2 классы множества состояний C .

Объекты классификации представляют собой кортеж значений $X = (x_1, \dots, x_n)$ длины $n \geq 2$.

Имеется некоторая начальная размеченная выборка $X \times L$. На обучающей выборке определяются известные пары $(X_b l_j)$.

Таким образом для входного вектора значений X необходимо построить алгоритм классификации, который по входному кортежу X отображает состояние Z во множество C классов событий: $a: X \times L \rightarrow C$.

Вывод

В первой главе работы было проведено обзорное исследование по теме идентификации состояния бортовых систем судна, была изучена теоретическая составляющая базовых терминов и определений понятия бортовая система и её устройств. Установлено, что бортовая система — это сложная составляющая любого судна независимо от его назначения. Кроме того, было установлено, что устройства бортовых систем требуют внимания и контроля за собой.

За этим следует, что задача идентификации состояния бортовых систем является сложной и актуальной, так как позволяет своевременно выявлять

неисправности, предотвращать аварийные ситуации и повышать безопасность судоходства.

ГЛАВА 2. ПРИМЕНЕНИЕ БЭГГИНГА В МОДЕЛЯХ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ИДЕНТИФИКАЦИИ СОСТОЯНИЯ БОРТОВЫХ СИСТЕМ

В данной главе рассматривается применение методов машинного обучения для идентификации состояния бортовых систем. В качестве основных методов машинного обучения используются алгоритмы классификации, регрессии и кластеризации, искусственные нейронные сети. Сферы, в которых применяют машинное обучение, рассмотрены. Отдельно описано про применение МО в морских информационных системах. Кроме того, в данном разделе упомянуто про примеры применения бэггинга в различных методах машинного обучения, таких как случайный лес, бустинг и другие. Таким образом, данная глава является важным шагом в разработке эффективных методов идентификации состояния устройств бортовых систем, что может улучшить безопасность плавания и экономическую эффективность работы судна.

2.1 Методы машинного обучения и способы повышения качества обработки информации для оценки состояния бортовых систем

Методы машинного обучения могут использоваться для анализа данных, собранных с бортовых систем. Эти методы могут помочь выявить аномалии и предсказать возможные сбои в работе системы.

Важным способом повышения качества обработки информации является улучшение качества сбора данных. Для этого могут использоваться различные датчики и дополнительные средства сбора информации.

Также можно использовать методы статистического анализа для проверки корректности данных и выявления аномалий. Это может помочь улучшить точность предсказаний и предотвратить возможные сбои в работе системы.

Важно также разработать эффективные алгоритмы обработки данных, которые будут учитывать специфику работы бортовых систем и позволят быстро и точно обрабатывать большие объемы информации.

Машинное обучение – это область искусственного интеллекта, которая изучает, как компьютеры могут обучаться на основе опыта и данных, чтобы решать задачи без явного программирования. Машинное обучение находит свое применение во многих областях, включая:

1. Медицину и биологию – использование машинного обучения для диагностики болезней, анализа медицинских изображений, прогнозирования эпидемий и разработки новых лекарств.

2. Финансы – использование машинного обучения для анализа финансовых данных, прогнозирования курсов валют, определения рисков и своевременного выявления мошенничества.

3. Транспорт – использование машинного обучения для управления автономными транспортными средствами, оптимизации маршрутов и управления трафиком.

4. Обработка естественного языка – использование машинного обучения для анализа и обработки естественных языков, создания персональных ассистентов и улучшения качества машинного перевода.

5. Производство – использование машинного обучения для управления производственными процессами, оптимизации производственных линий и прогнозирования спроса на продукцию.

6. Реклама – использование машинного обучения для создания персонализированной рекламы, прогнозирования поведения потребителей и оптимизации рекламной кампании.

7. Обработка изображений – использование машинного обучения для распознавания лиц, анализа изображений, детектирования объектов и улучшения качества фотографий.

8. Интернет вещей – использование машинного обучения для анализа данных, собранных с датчиков и устройств IoT (Internet of Things), для

улучшения управления энергопотреблением, оптимизации производственных процессов и улучшения качества жизни людей.

Машинное обучение имеет огромный потенциал и может быть применено во многих других областях.

Кроме того, в области морских информационных систем машинное обучение используется для решения различных задач, связанных с обработкой и анализом данных, полученных в процессе наблюдения за состоянием морской среды и ее изменениями. Например, с помощью машинного обучения можно:

- Анализировать данные, полученные с помощью датчиков и приборов, установленных на судах и буях, для выявления изменений в состоянии морской среды и прогнозирования ее развития.

- Разрабатывать модели, позволяющие определять оптимальные маршруты для судов, учитывая различные факторы, такие как погода, течения, глубина моря и другие.

- Анализировать данные, полученные с помощью спутникового мониторинга, для выявления изменений в состоянии окружающей среды, таких как изменение температуры и химического состава воды, наличие водорослей и т.д.

- Разрабатывать системы автоматического управления и наблюдения за морскими объектами, такими как подводные аппараты и беспилотные суда.

Для решения этих и других задач используются различные методы машинного обучения, такие как нейронные сети, методы кластеризации и классификации, анализ временных рядов и другие.

Метод обучения с учителем (Supervised Learning) - это один из основных подходов в машинном обучении, при котором модель обучается на основе маркированных данных, где каждый пример данных имеет известную метку или ответ, которые модель должна предсказывать. Этот метод заключается в том, что алгоритму предоставляются пары входных и выходных данных (обучающие примеры) и на основе этих данных алгоритм

строит модель, которая может предсказывать результаты на новых данных. Примеры задач, решаемых методом обучения с учителем, включают классификацию, регрессию и прогнозирование.

Процесс обучения с учителем включает в себя следующие шаги:

1. Подготовка данных: данные разбиваются на обучающую и тестовую выборки. Обучающая выборка используется для обучения модели, а тестовая - для оценки ее качества.

2. Выбор модели: выбирается модель, которая будет использоваться для обучения. Модель зависит от типа задачи, например, для задач классификации может использоваться логистическая регрессия или случайный лес, а для задач регрессии - линейная регрессия или нейронные сети.

3. Обучение модели: модель обучается на обучающей выборке с использованием алгоритма обучения, например, градиентного спуска или метода опорных векторов.

4. Оценка качества модели: после обучения модели оценивается ее качество на тестовой выборке с помощью метрик, таких как точность, полнота и F1-мера для задач классификации или среднеквадратическая ошибка для задач регрессии.

5. Использование модели: после успешной оценки качества модели, она может быть использована для предсказания ответов на новых данных.

Метод обучения с учителем является одним из наиболее широко используемых методов в машинном обучении, и он применяется в различных областях, таких как анализ данных, компьютерное зрение, речевые технологии, биоинформатика и другие.

Байесовский метод машинного обучения [14] основан на теореме Байеса, которая позволяет обновлять вероятностную модель на основе новых данных. Этот метод используется для оценки вероятностей классов и моделей на основе имеющихся данных. В байесовском методе машинного обучения сначала определяется априорное распределение вероятностей

классов и моделей, которое может быть задано экспертами или на основе предыдущих данных. Затем, при поступлении новых данных, обновляется апостериорное распределение вероятностей, которое учитывает как априорную информацию, так и новые данные. Байесовский метод машинного обучения позволяет учитывать неопределенность и шум в данных, а также позволяет решать задачи классификации, регрессии, кластеризации и другие. Он может быть использован как самостоятельный метод, так и в сочетании с другими методами машинного обучения.

Пусть X множество объектов, Y конечное множество имён классов, множество $X \times Y$ является вероятностным пространством с плотностью распределения $p(x, y) = P(y)p(x|y)$. Вероятности появления объектов каждого из классов $P_y = P(y)$ называются априорными вероятностями классов. Плотности распределения $p_y(x) = p(x|y)$ называются функциями правдоподобия классов.

Функционалом среднего риска называется ожидаемая величина потери при классификации объектов алгоритмом a . Записывается он по формуле (1).

$$R(a) = \sum_{y \in Y} \sum_{s \in Y} \lambda_{ys} P_y P(A_s|y). \quad (1)$$

где λ_{ys} - величина потери

$P_y P(A_s|y)$ - вероятность того, что появится объект класса y и алгоритм a отнесет его к классу s .

Теорема об оптимальности байесовского классификатора:

Если известны априорные вероятности P_y и функции правдоподобия $p_y(x)$, то минимум среднего риска $R(a)$ достигается алгоритмом и мы получаем выражение (2).

$$a(x) = \arg \min_{y \in Y} \sum \lambda_{ys} P_y p_y(x) \quad (2)$$

где $a(x)$ - алгоритм, который нужно построить

P_y - априорная вероятность

$p_y(x)$ - плотность распределения

Метрические методы классификации [15] основаны на измерении расстояний между объектами в пространстве признаков. Эти методы используют метрику, которая определяет, как расстояние между двумя объектами в пространстве признаков определяется.

Одним из наиболее распространенных метрических методов является метод ближайших соседей (k-NN), который определяет класс нового объекта, основываясь на классах его ближайших соседей.

Другими метрическими методами являются метод парзеновского окна, метод потенциальных функций и метод минимального расстояния.

Метрические методы классификации имеют ряд преимуществ, таких как простота реализации и высокая точность при условии правильного выбора метрики. Однако, они также имеют свои недостатки, такие как чувствительность к выбросам и сложность выбора оптимального значения k в методе ближайших соседей. Смотреть рис.1.

Алгоритм k ближайших соседей [15] (k nearest neighbors, k NN). Чтобы сгладить влияние выбросов, будем относить объект u к тому классу, элементов которого окажется больше среди k ближайших соседей $x_u^{(i)}, i = 1, \dots, k$, отсюда получается формула (3).

$$a(u; X^l, k) = \operatorname{argmax}_{y \in Y} \sum_{i=1}^k [y_u^{(i)} = y] \quad (3)$$

где $a(u; X^l, k)$ - алгоритм, который нужно построить

При $k = 1$ этот алгоритм совпадает с предыдущим, следовательно, неустойчив к шуму. При $k = 1$, наоборот, он чрезмерно устойчив и вырождается в константу.

Таким образом, крайние значения k нежелательны. На практике оптимальное значение параметра k определяют по критерию скользящего контроля с исключением объектов по одному (leave-one-out, LOO). Для каждого объекта $x_i \in X^l$ проверяется, правильно ли он классифицируется по своим k ближайшим соседям.

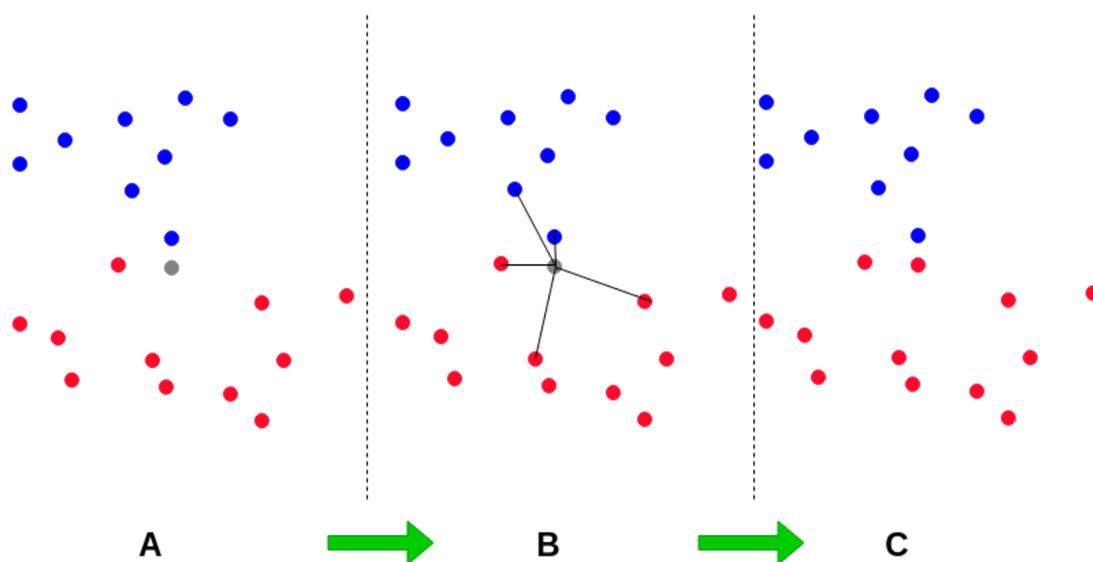


Рис.1. Пример классификации, методом 5 ближайших соседей

Линейные методы классификации [16] заключаются в том, что они строят линейную границу между классами в многомерном пространстве признаков. Это означает, что признаки умножаются на весовые коэффициенты и складываются, чтобы получить оценку класса. Если оценка превышает пороговое значение, то объект относится к одному классу, иначе – к другому. Линейные методы классификации часто используются для решения задач, где классы хорошо разделены гиперплоскостью, например, в задачах распознавания образов. Гипотеза может быть представлена как линейная комбинация входных признаков с весами, которые определяют вклад каждого признака в окончательное решение.

Рассмотрим метод опорных векторов. Метод опорных векторов (Support Vector Machines, SVM) – это алгоритм машинного обучения, который используется для решения задач классификации и регрессии. Он основан на поиске гиперплоскости в многомерном пространстве, которая разделяет два класса объектов.

Основной идеей метода является поиск такой гиперплоскости, которая максимально удалена от ближайших объектов двух классов. Эти объекты называются опорными векторами. Гиперплоскость, найденная методом, позволяет классифицировать новые объекты в зависимости от того, на какой стороне гиперплоскости они находятся.

SVM имеет несколько преимуществ перед другими алгоритмами машинного обучения. Во-первых, он позволяет решать задачи с большим количеством признаков и объектов. Во-вторых, SVM применим для разных типов данных, включая текстовые, изображения и числовые. В-третьих, этот метод позволяет решать задачи с нелинейными зависимостями между признаками.

Недостатками метода являются его вычислительная сложность и чувствительность к выбросам в данных. Также выбор правильного ядра (функции, определяющей кривизну гиперплоскости) может оказаться нетривиальной задачей.

В целом, SVM является мощным инструментом машинного обучения, который может быть использован для решения многих задач классификации и регрессии.

Линейно неразделимая выборка [16]. Рассмотрим задачу классификации на два непересекающихся класса, в которой объекты описываются n -мерными вещественными векторами: $X = R^n, Y = \{-1, +1\}$.

Будем строить линейный пороговый классификатор (4):

$$a(x) = \text{sign}\left(\sum_{j=1}^n \omega_j x^j - \omega_0\right) = \text{sign}(\langle \omega, x \rangle - \omega_0) \quad (4)$$

где $x = (x^1, \dots, x^n)$ – признаковое описание объекта x
 вектор $\omega = (\omega^1, \dots, \omega^n) \in R^n$ и скалярный порог $\omega_0 \in R$ являются параметрами алгоритма

уравнение $\langle \omega, x \rangle = \omega_0$ описывает гиперплоскость, разделяющую классы в пространстве R^n .

В случае линейно разделимой выборки получаем задачу квадратичного программирования (5) : требуется найти значения параметров ω и ω_0 , при которых выполняются l ограничений-неравенств и норма вектора ω минимальна:

$$\begin{cases} \langle \omega, \omega \rangle \rightarrow \min; \\ y_i (\langle \omega, x_i \rangle - \omega_0) \geq 1, \quad i = 1, \dots, l. \end{cases} \quad (5)$$

Где вектор $\omega = (\omega^1, \dots, \omega^n) \in R^n$ и скалярный порог $\omega_0 \in R$ являются параметрами алгоритма

уравнение $\langle \omega, x \rangle = \omega_0$ описывает гиперплоскость, разделяющую классы в пространстве R^n .

На практике линейно разделимые выборки [16] практически не встречаются: в данных возможны выбросы и нечеткие границы между классами. В таком случае поставленная выше задача не имеет решений, и необходимо ослабить ограничения, позволив некоторым объектам попадать на «территорию» другого класса. Для каждого объекта отнимем от отступа некоторую положительную величину ξ_i , но потребуем чтобы эти введенные поправки были минимальны. Это приведёт к следующей постановке задачи, называемой также *SVM* с мягким отступом, формула (6).

$$\frac{1}{2} \|\varpi\|^2 + C \sum_{i=1}^l \xi_i \rightarrow \min$$

$$M_i(\varpi, b) \geq 1 - \xi_i, \quad i = 1, \dots, l$$

$$\xi_i \geq 0, \quad i = 1, \dots, l$$
(6)

где ξ_i , - некоторая положительная величина, которую отнимают от отступа

Методы восстановления регрессии.

Методы восстановления регрессии [17] - это методы, которые используются для построения моделей регрессии на основе имеющихся данных. Они заключаются в том, чтобы найти математическую связь между зависимой переменной и одной или несколькими независимыми переменными.

Рассмотрим метод наименьших квадратов - это метод, который минимизирует сумму квадратов расстояний между точками данных и линией регрессии. Этот метод является наиболее распространенным и часто используется для построения моделей линейной регрессии.

Пусть задана модель регрессии — параметрическое семейство функций $g(x, a)$, где $a \in R^p$ — вектор параметров модели. Определим функционал качества аппроксимации целевой зависимости на выборке X^l как сумму квадратов ошибок (7);

$$Q(a, X^l) = \sum_{i=1}^l (g(x_i, a) - y_i)^2.$$
(7)

где $g(x, a)$ - параметрическое семейство функций

$a \in R^p$ — вектор параметров модели

X^l - выборка

Обучение по методу наименьших квадратов (МНК) состоит в том, чтобы найти вектор параметров a^* , при котором достигается минимум среднего квадрата ошибки на заданной обучающей выборке X^l :

$$a^* = \arg \min_{a \in R^p} Q(a, X^l). \quad (8)$$

где a^* - вектор параметров

X^l – обучающая выборка

Стандартный способ решения этой оптимизационной задачи — воспользоваться необходимым условием минимума (9). Если функция $g(x, a)$ достаточное число раз дифференцируема по a , то в точке минимума выполняется система p уравнений относительно p неизвестных:

$$\frac{\partial Q}{\partial a}(a, X^l) = 2 \sum_{i=1}^l (g(x_i, a) - y_i) \frac{\partial g}{\partial a}(x_i, a) = 0 \quad (9)$$

где a^* - вектор параметров

X^l – обучающая выборка

$g(x, a)$ - параметрическое семейство функций

Искусственные нейронные сети (ИНС) [17] - это модели, используемые в машинном обучении, которые имитируют работу человеческого мозга. Они состоят из множества связанных между собой "нейронов", которые обрабатывают входные данные и выдают результат в виде выходных значений.

ИНС используются в различных областях, включая компьютерное зрение, распознавание речи, обработку естественных языков и другие. Они могут использоваться для решения задач классификации, регрессии и кластеризации.

ИНС обучаются на основе данных, используя алгоритмы обратного распространения ошибки. В процессе обучения ИНС корректируют веса связей между нейронами, чтобы минимизировать ошибку предсказания.

Одним из преимуществ ИНС является их способность обрабатывать большие объемы данных и выявлять скрытые закономерности. Однако, они могут быть чувствительны к переобучению, если обучение проводится на недостаточно разнообразных данных.

В целом, ИНС являются мощным инструментом в машинном обучении и могут быть использованы для решения многих задач, но их эффективность зависит от правильного выбора архитектуры и параметров обучения.

Рассмотрим метод обратного распространения ошибок, метод обратного распространения ошибок (backpropagation) является одним из наиболее популярных алгоритмов обучения нейронных сетей. Этот метод используется для обновления весовых коэффициентов в сети на основе полученной ошибки.

Процесс обучения нейронной сети с использованием метода обратного распространения ошибок начинается с передачи входного сигнала через слои нейронов. Каждый слой рассчитывает выходное значение на основе входного сигнала и весовых коэффициентов.

Затем выходное значение сравнивается с ожидаемым значением, и вычисляется ошибка сети. Эта ошибка распространяется обратно через слои нейронов, начиная с последнего слоя и двигаясь к первому слою. Каждый слой получает ошибку от следующего слоя и вычисляет свою собственную ошибку на основе этой информации.

Затем весовые коэффициенты сети обновляются на основе вычисленной ошибки. Более тщательно распространяя ошибку обратно через слои, метод обратного распространения ошибок помогает нейронной сети "выучиться" и настроить свои весовые коэффициенты для более точного прогнозирования выходных значений на основе входных данных.

Кластеризация - это процесс разделения множества данных на группы (кластеры), которые имеют схожие характеристики. Кластеры могут быть созданы на основе различных критериев, таких как расстояние между точками, сходство признаков и т.д.

Визуализация - это процесс представления данных в графическом виде, чтобы упростить их понимание. Визуализация может быть полезна для обнаружения паттернов и трендов в данных, а также для сравнения различных наборов данных.

Кластеризация и визуализация могут быть использованы вместе для анализа данных. Например, после кластеризации данных, можно построить диаграмму рассеяния, чтобы визуализировать распределение кластеров в пространстве. Это может помочь понять структуру данных и выявить скрытые закономерности.

Кроме того, визуализация может быть полезна для интерпретации результатов кластеризации, особенно когда кластеры имеют многомерную структуру. Визуализация может помочь идентифицировать наиболее значимые признаки, которые разделяют кластеры, и помочь в дальнейшем анализе данных.

Ансамблевые методы машинного обучения [18] - это методы, которые используют несколько моделей машинного обучения для улучшения качества предсказаний. Они работают путем комбинирования нескольких моделей, каждая из которых может быть обучена на разных подмножествах обучающих данных или с использованием разных алгоритмов.

Существуют различные типы ансамблевых методов машинного обучения, включая:

1. Бэггинг (Bootstrap Aggregating) - метод, который использует несколько моделей, каждая из которых обучается на случайном подмножестве обучающих данных [1]. На рисунке 2 схематично изображено обучение мета-модели.

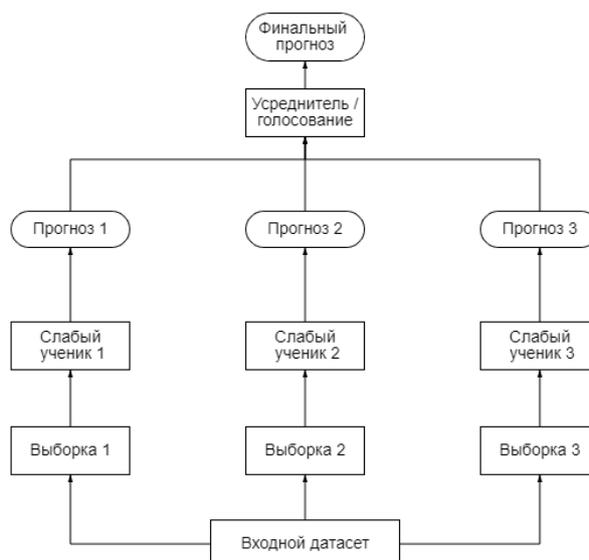


Рис.2. Обучение мета-модели беггингом

2. Бустинг (Boosting). Метод бустинга подобен методу бэггинга в том, что в обоих случаях используется множество одинаковых моделей, объединенных для создания более сильного ученика. Однако, отличие заключается в том, что модели в бустинге приспособливаются к данным последовательно, исправляя ошибки предыдущих моделей.

Для бустинга выбираются базовые модели с низким разбросом и высоким смещением, такие как неглубокие деревья решений. Это связано с тем, что такие модели требуют меньше вычислительных ресурсов. Кроме того, бустинг не может быть распараллелен.

Существуют два основных алгоритма бустинга - адаптивный и градиентный.

3. Стекинг (Stacking) - метод, который использует несколько моделей, каждая из которых обучается на обучающих данных и предсказывает результаты для тестовых данных. Затем эти результаты используются как входные данные для более высокоуровневой модели, которая делает окончательное предсказание. На рисунке 3 схематично изображено обучение мета-модели.

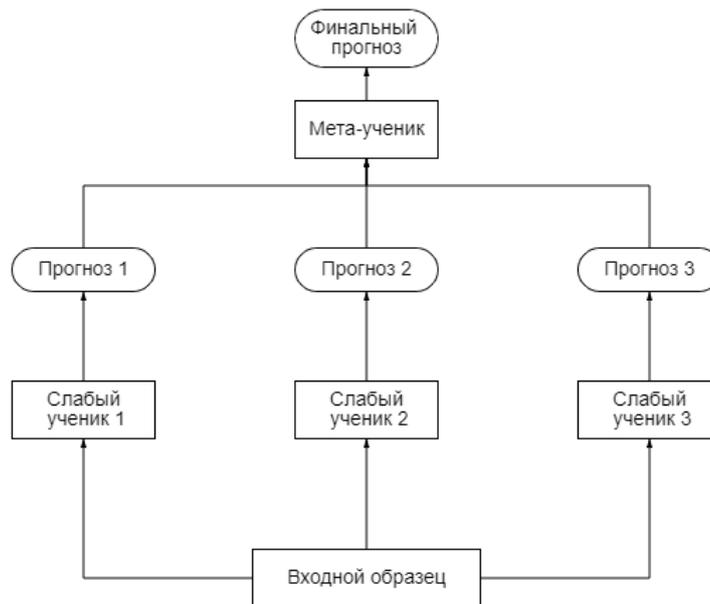


Рис.3. Обучение мета-модели стекингом

Ансамблевые методы машинного обучения обычно дают лучшие результаты, чем отдельные модели, поскольку они уменьшают ошибку предсказания, повышают устойчивость модели к шуму и улучшают обобщающую способность модели.

2.2 Применение бэггинга в методах МО для повышения качества оценки состояния бортовых систем

Метод беггинга (bootstrap aggregating) – это один из наиболее популярных методов ансамблевого машинного обучения, который позволяет улучшить качество классификации или регрессии путем комбинирования нескольких моделей. [1]

Суть метода заключается в использовании случайных выборок из обучающего набора данных для обучения отдельных моделей, а затем комбинировании результатов их работы. Таким образом, каждая модель обучается на своей подвыборке данных, что позволяет уменьшить влияние шума и выбросов в исходном наборе данных и повысить обобщающую способность модели.

Процесс обучения метода беггинга состоит из следующих шагов:

1. Создание случайных подвыборок набора данных. Для этого из исходного набора данных выбирается случайная выборка с повторениями, размер которой равен размеру исходного набора данных.

2. Обучение отдельных моделей на каждой подвыборке данных. Для этого на каждой подвыборке данных обучается отдельная модель с использованием выбранного алгоритма машинного обучения.

3. Комбинирование результатов работы отдельных моделей. Для получения окончательного результата работы алгоритма совмещаются результаты работы отдельных моделей. В случае классификации, для принятия решения о классификации объекта используется голосование большинства моделей. В случае регрессии, для прогнозирования значения используется усреднение значений, предсказанных отдельными моделями.

Преимущества метода беггинга:

- Уменьшение влияния шума и выбросов в исходном наборе данных.
- Уменьшение переобучения модели.
- Повышение обобщающей способности модели.
- Увеличение точности предсказаний.

Недостатки метода беггинга:

- Требуется большое количество вычислительных ресурсов для обучения нескольких моделей.
- Не всегда эффективен в случае использования сложных моделей машинного обучения.
- Может привести к увеличению ошибки в случае использования несовместимых моделей.

Идея алгоритма очень проста и была предложена Лео Брейманом в 1994 году. Этот подход оказался весьма эффективным и используется до сих пор. Итак, предположим изначально у нас имеется некоторая обучающая выборка: $X^l = \{(x_i, y_i)_{i=1}^l\}$

На ее основе мы хотим сформировать несколько разных и, в общем случае, независимых алгоритмов обработки входного вектора x : $a_1(x), a_2(x), \dots, a_T(x)$

А, затем, усреднить ответы каждого алгоритма для формирования общего решения.

Идея, когда мы формируем множество независимых алгоритмов, каждый выдает свой вариант ответа, а затем, мы его усредняем, чтобы получить более точное значение заложена в бэггинге.

Так как же нам сформировать T независимых алгоритмов, используя всего одну обучающую выборку? Здесь есть несколько идей, но в бэггинге используется очень простой подход, который носит название **бутстрэп** (bootstrap). Суть бутстрэпа заключается в формировании T новых обучающих выборок на основе одной исходной $X^l = \{(x_i, y_i)_{i=1}^l\}$. Для этого случайным образом выбирается k -й элемент (x_k, y_k) выборки и копируется в новую (в прежней он остается, не удаляется). Затем, эта операция повторяется t раз – по заданному размеру новой выборки. Так формируется новая обучающая выборка, состоящая из элементов исходной с некоторыми повторениями, так как вполне можно несколько раз случайно отобрать один и тот же элемент. После формирования одной выборки, переходят к формированию следующей и так T раз для T выборок, которые, очевидно, будут несколько отличаться друг от друга. Это идея бутстрэпа (см. рис.4).

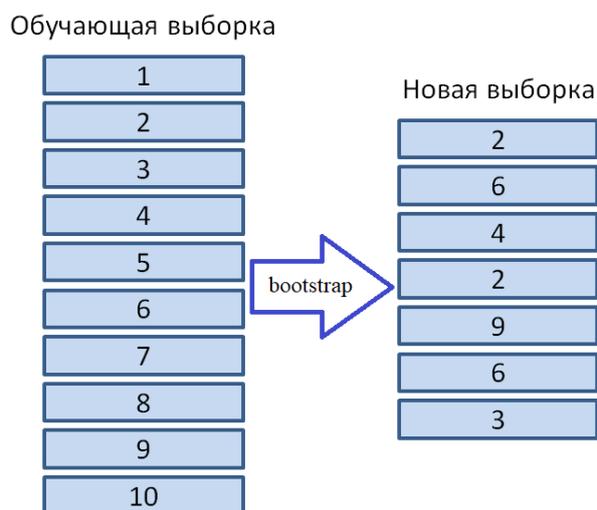


Рис.4. Образование новой выборки

Можно показать, что формируя, таким образом, новую обучающую выборку длиной l , она будет использовать, в среднем, $(1 - \frac{1}{e}) \cdot 100\% \approx 63,2\%$ образов из исходной выборки (остальные будут повторяться). Это значит, что оставшаяся часть выборки в $36,8\%$ можно использовать в качестве отложенной для проверки качества алгоритма (10). И это нашло выражение в критерии под названием *out-of bag*, когда мы определяем число ошибок по объектам x_i , не участвующих в обучении того или иного дерева:

$$out - of - bag(a) = \sum_{i=1}^T \left[sign\left(\sum_{i=1}^T [x_i \notin U_t] b_t(x_i)\right) \neq y_t \right] \rightarrow min \quad (10)$$

где $U_t, t = 1, \dots, T$ - множества объектов, составляющих обучающую выборку для дерева t .

out-of bag – это несмещенная оценка обобщающей способности итогового алгоритма $a(x)$.

Бэггинг с решающими деревьями. Случайный лес.

Собственно, само слово *bagging* произошло от сокращения двух английских слов: *bootstrap aggregation*. Итак, имея T случайных выборок

длиной m элементов ($m < l$), мы можем по ним получить T алгоритмов (классификации или регрессии). Например, используя линейную модель: $a(x) = \langle \omega, x \rangle = \omega^T \cdot x$ можно сформировать T наборов весовых векторов: $\omega_1, \omega_2, \dots, \omega_T$ каждая для своей обучающей выборки. В результате получаем T различных алгоритмов: $a_j(x) = \langle \omega_j, x \rangle$, $j = 1, 2, \dots, T$

А, затем, усредняя ответы от них, формируем общий результат: $a(x) = \frac{1}{T} \sum_{j=1}^T a_j(x)$. Или, устраивая голосование, решаем задачу классификации: $a(x) = \text{sign}(\frac{1}{T} \sum_{j=1}^T a_j(x))$ алгоритмы в своей совокупности должны охватывать как можно больше возможных исходов для каждого входного вектора x и, кроме того, формировать как можно более независимые ответы. В этом смысле обычные линейные модели не очень пригодны при композиции (в частности, при усреднении ответов). Гораздо лучшие результаты дают решающие деревья, построенные независимо на каждой сформированной обучающей выборке (на этапе бустрэпа).

Как ни странно, приверженность деревьев к переобучению играет положительную роль. Они, во-первых, получаются довольно разнообразными и, во-вторых, описывают самые разные исходы для входных векторов x . Затем, при усреднении результатов, эффект переобучения естественным образом нивелируется (уменьшается) и итоговое выходное значение оказывается достаточно точным и устойчивым к отдельным выбросам. В ряде задач точность оказывается выше всех других подходов машинного обучения. Именно поэтому бутстрэп быстро завоевал свою популярность.

Чтобы решающие деревья получались еще более разнообразными и формировали менее зависимые ответы, предлагается при их обучении в каждой промежуточной вершине случайным образом отбирать некоторое количество признаков $m < n$ и уже среди них отбирать лучшие для ветвления. Наборы из таких деревьев называют случайным лесом (random forest). Причем, было показано, что в задачах классификации число $m =$

$\lceil \sqrt{n} \rceil$, а в задачах регрессии $m = \lceil \frac{n}{3} \rceil$. Есть теоретические выкладки почему это так. Сами же признаки и пороги для ветвления выбираются, как правило, по критерию Джини (он быстрее вычисляется, чем энтропийный и приводит к практически тем же результатам). А усечений деревьев уже не делают, они остаются переобученными. Как отмечалось ранее, обобщение и устойчивость выходного значения будет определяться усреднением независимых ответов от каждого дерева. Причем, для случайного леса кривые качества на обучающей и проверочной выборках в среднем уменьшаются при увеличении числа деревьев T рисунок 5:

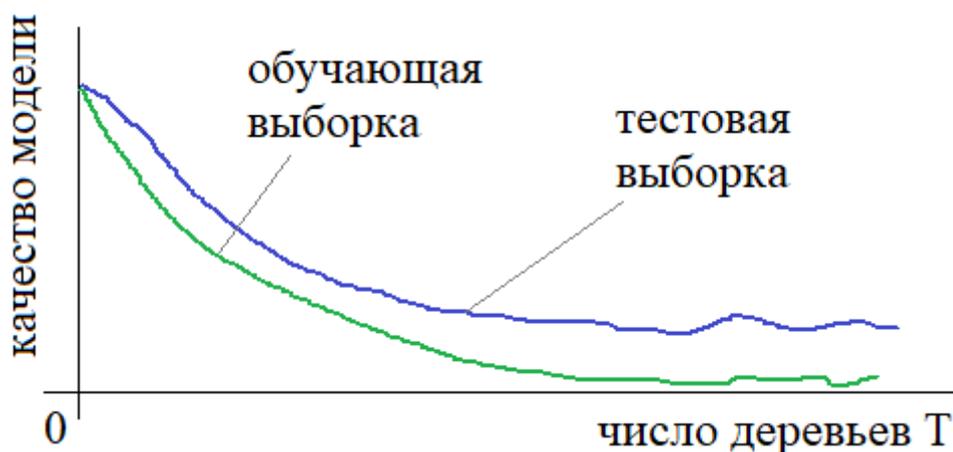


Рис. 5. Графики обучающей и тестовой выборок

То есть, с ростом числа T решающих деревьев в случайном лесе итоговая модель не переобучается, а лишь достигает некоторого предельного уровня качества. Это очень хорошее свойство случайного леса, так как мы, фактически, получаем алгоритм, в котором нет сложно настраиваемых гиперпараметров. Единственный параметр T можно взять, например, 100, а затем, 500 и сравнить результаты на тестовой выборке. То есть, подобрать его очень просто. Остальные гиперпараметры для построения решающих деревьев можно выбирать с позиции некоего здравого смысла. Главное, чтобы деревья получались глубокими, гарантируя малые смещения в ответах.

Напомним, что под смещением здесь понимается стремление к нулю (в среднем) ошибки прогноза: $E\{(x - a(x))\} \rightarrow 0$

Как и любой алгоритм, случайные леса имеют свои преимущества и недостатки. К преимуществам можно отнести:

- имеет высокую точность прогнозов (на большинстве задач работает лучше линейных алгоритмов); точность сравнима с точностью бустинга;
- практически не чувствителен к выбросам в данных из-за случайного сэмплирования выборок методом бутстрэпа;
- не чувствителен к масштабированию (и вообще к любым монотонным преобразованиям) значений признаков, связано с выбором случайных подпространств;
- не требует тщательной настройки параметров, хорошо работает «из коробки»;
- способен эффективно обрабатывать данные с большим числом признаков и классов;
- редко переобучается, на практике добавление деревьев почти всегда только улучшает композицию (до определенного, предельного уровня);
- хорошо работает с пропущенными данными; сохраняет хорошую точность, если большая часть данных пропущена;
- могут быть расширены до неразмеченных данных, что приводит к возможности делать кластеризацию и визуализацию данных, обнаруживать выбросы;
- легко распараллеливать и масштабировать (увеличивать число деревьев и их глубину).

Недостатки случайного леса

- в отличие от одного дерева, результаты случайного леса сложнее интерпретировать;

- алгоритм работает хуже многих линейных методов, когда в выборке очень много разреженных признаков (тексты, Bag of words);
- случайный лес не умеет экстраполировать, в отличие от той же линейной регрессии;
- алгоритм склонен к переобучению на некоторых задачах, особенно на сильно зашумленных данных;
- для данных, включающих категориальные переменные с различным количеством уровней, случайные леса предвзяты в пользу признаков с большим количеством уровней: когда у признака много уровней, дерево будет сильнее подстраиваться именно под эти признаки, так как на них можно получить более высокое значение оптимизируемого функционала (информационный выигрыш);
- большой размер получающихся моделей. Требуется $O(N \cdot K)$ памяти для хранения модели, где K – число деревьев.

ГЛАВА 3. АНАЛИЗ ПРИМЕНИМОСТИ ПРЕДЛАГАЕМОГО ПОДХОДА

3.1 Описание предлагаемого подхода

В процессе разработки устройств бортовых систем водного судна используются различные методы и технологии, такие как компьютерное моделирование, тестирование в экстремальных условиях, анализ данных, применение современных материалов и технологий производства, тем самым усложняет анализ исходного кода. Методы быстрой разработки программных и аппаратных частей, использующие стандартные компоненты от разных производителей, приводят к созданию устройств, которые могут быть описаны как «черный ящик».

Бортовые узлы имеют ограниченные вычислительные ресурсы и способны выполнять лишь небольшой набор команд, что позволяет обрабатывать и идентифицировать ограниченное количество состояний и переходов между ними.

Во время функционирования процессы устройств бортовых систем протекают в динамике, одновременно меняется множество параметров.

Состояние внешней среды $u(t)$, вызванное поступлением на устройство команд управления, приемом, передачей сообщений, функционирование элемента, определяемое внутренними ситуациями обработки данных и реализации вычислительных алгоритмов, характеризующееся переходными характеристики $h(t)$ дает возможность рассмотреть устройство как динамическую систему. Имеется q входов и d выходов [8], на вход подается управляющая команда и значения переменных внешней среды, на выходе появляются сигналы $S(t)$ (например, показывающие загрузку ресурсов), регистрируемые различными датчиками. Получаемые по внешним каналам значения сигналов содержат значения шумовой составляющей $v(t)v(t)$,

определяемой свойствами измерительного прибора, характеристик получаемого сигнала и т.д.

Модель состояния устройства бортовой системы определяется соотношением [3]:

$$\sum_{i=1}^q \sum_{j=1}^d \int_0^t u_i(t) h_{ij}(t-\tau) d\tau = \sum_{j=1}^d \int_0^t f(s_j(t-\tau), v_j(t-\tau)) d\tau \quad (11)$$

где q – количество каналов источников;

h_{ij} – переходные характеристики i -го канала для jj -го регистрирующего, получаемые по каналу значения датчика;

f – функция измеренных значений.

В дискретные моменты времени функционирования устройства t_0, t_1, \dots, t_n происходит регистрация векторов числовых последовательностей. Значения $X(t)$ отражают данные, полученные от датчиков, содержащие смесь полезного сигнала $S(t)$ и шума, выраженного параметром $v(t)$:

$$X(t) = F[S(t), v(t)]. \quad (12)$$

где вектор X является результатом смешанных, взаимно независимых сигналов $S(t)$, имеющих искажение шумовой составляющей $v(t)$. Вектор X представляет собой временной ряд значений, полученный от регистрирующих устройств.

Векторы X_1, X_2, \dots, X_n отражают поведение процесса в многомерном координатном пространстве и определяют множество состояний Z . Состояния разделяются множеством классов C , где подмножества делятся на опасные C_1 и безопасные C_2 состояния.

Таким образом, имеется размеченная конечная обучающая выборка $X = \{(x_{11}, \dots, x_{n1}), (x_{12}, \dots, x_{n2}), \dots, (x_{1m}, \dots, x_{nm}), \dots\}$:

Требуется построить алгоритм классификации a входного вектора X_i для отображения $Z \rightarrow C$.

Размеченная обучающая выборка содержит значения временных рядов от регистрирующих устройств в заранее определенных состояниях и режимах работы. Известные состояния $\{z_1, \dots, z_l\} \in Z$, определены только на объектах наблюдаемых последовательностей $\{(x_{11}, \dots, x_{n1}), (x_{12}, \dots, x_{n2}), \dots, (x_{1m}, \dots, x_{nm})\}$.

От исследуемого устройства бортового узла на интервале $t_0 \leq t \leq T$ наблюдается случайная векторная функция $X(t) = f(S(t), v(t))$, где в дискретные моменты времени t_0, t_1, \dots, t_k , регистрируется временной ряд $x_i = X(t_i)$.

Определено множество классов состояний $C = \{c_0, c_1, \dots, c_n\}$, в одном из которых в дискретный момент времени t_j может находиться система.

Имеется k независимо друг от друга обученных классификаторов a_i , $i=1, \dots, k$. X – множество наборов признаков. $a_i(x_i) \rightarrow_j \in C$ – ответ i -го классификатора. $\{P_i(c_j|x_j)\}_j^n = 0$ – апостериорная вероятность для i -го классификатора после обучения. $w_i = \frac{1}{k}$ – весовые коэффициенты. $a(x) = \arg \max_{j=0, \dots, n} \sum_{i=0}^k w_i P_i(c_j|x_j)$ – общий классификатор.

Схема модели параллельной последовательности классификаторов приведена на рисунке 6. Подобные модели могут обучаться независимо друг от друга, что дает возможность осуществлять распараллеливание процессов. Предложенный подход идентификации состояния отличается использованием технологии классификации, реализующей композиции независимо обученных алгоритмов, обрабатывающих временные ряды, отражающих функционирование устройства во время выполнения процессов, что позволяет определять состояние устройства, не увеличивая объема хранящейся информации.

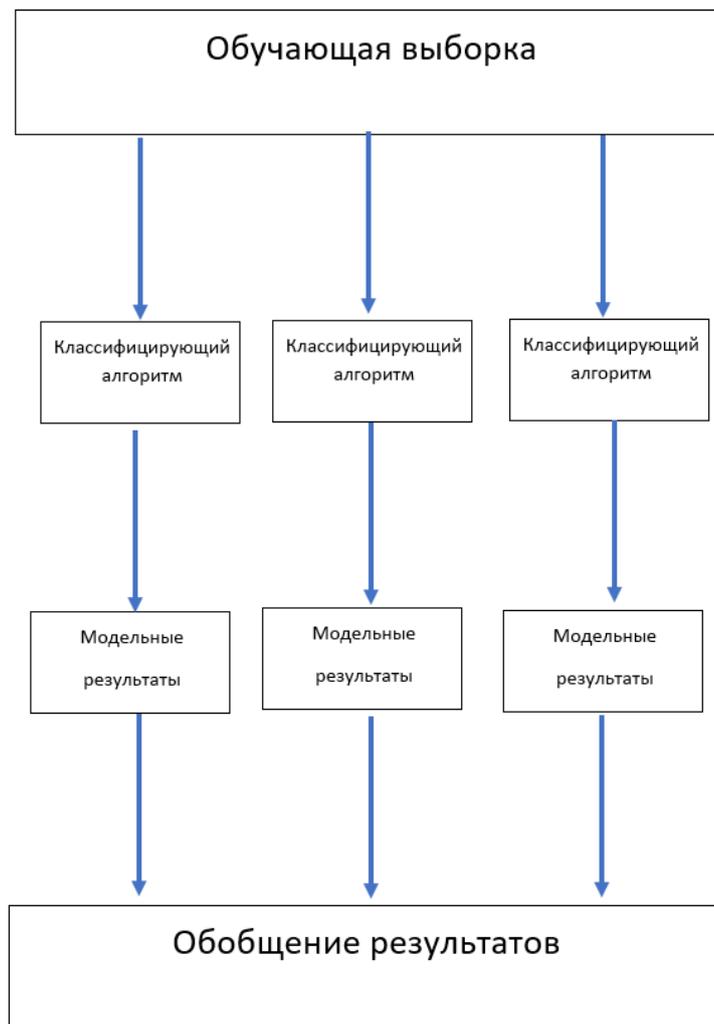


Рис.6. Схема модели параллельной последовательности классификаторов

Использование предлагаемого подхода на первоначальном этапе предполагает «настройку» устройства в заранее заданных режимах работы, где происходит предобработка на основе обучающей выборки.

3.2 Оценка применимости предлагаемого подхода

В рамках выпускной квалификационной работы для проведения эксперимента была использована установка «Крокус-3Т» ЕИЛЮ.464511.015. Аппаратура передачи данных (АПД) «Крокус-3Т» ЕИЛЮ.464511.015 предназначена для обмена дискретной информацией по некоммутируемым стандартным каналам тональной частоты (ТЧ), а также по широкополосным

каналам связи и физическим линиям (ФЛ), с одним или двумя абонентами, работающими по алгоритмам обмена информацией Аккорд СС-ПД, АККОРД СС-ПС и АРАГВА, а также по протоколу X25.2. На АПД реализованы разные скорости передачи данных, за счет изменения канала связи или метода модуляции, и могут быть равны 1200, 2400, 4800 и 9600 бит/с. Установка обеспечивает заданные условия обмена информацией как при эксплуатации на стационарных объектах, так и в составе изделий работающих на ходу. АПД «Крокус-3Т» серийно производится на предприятии СПбФ АО «НПК «ТРИСТАН», предназначена для передачи радиолокационной информации, входит в состав различных серийно-выпускаемых радиолокационных комплексов, применяемых в береговых радиотехнических частях военно-морского флота. Все сопутствующие нормативные документы были изучены в процессе прохождения производственной практики и написания выпускной квалификационной работы. Для оценки предлагаемого подхода на предприятии был развернут экспериментальный стенд, состоящий из АПД «Крокус-3Т», ВРМО «Ваенга-П», вспомогательного оборудования и программного обеспечения.

Оценка предлагаемого подхода осуществлялся на основе эксперимента, в ходе которого производилось выявление состояния, определяемого алгоритмом обработки данных, вычислительного узла. В качестве входных данных использовались временные ряды, отражающие загрузку вычислительных ресурсов, регистрируемые программой монитором. Схема эксперимента приведена на рис. 7.

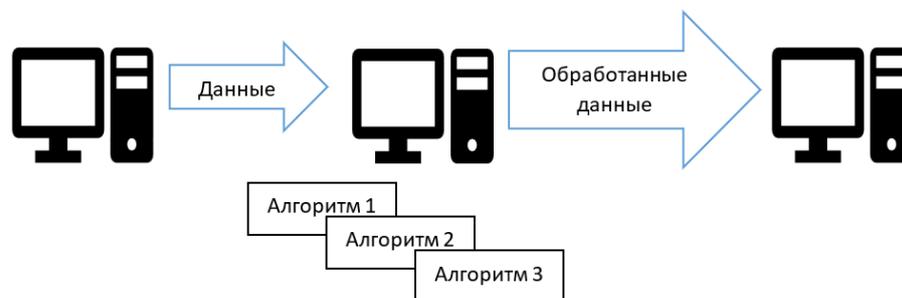


Рис. 7. Схема проведения эксперимента

На вычислительном устройстве запускались различные алгоритмы. В состоянии Z_1 функционировали только фоновые процессы. Во втором случае – узел C выступал в качестве транзитного узла, передававшего поступающую информацию без обработки (состояние Z_2). В третьей ситуации (состояние Z_3) кроме процессов приема и передачи дополнительно проводились процессы поиска заранее заданной информации (рис. 8-11).

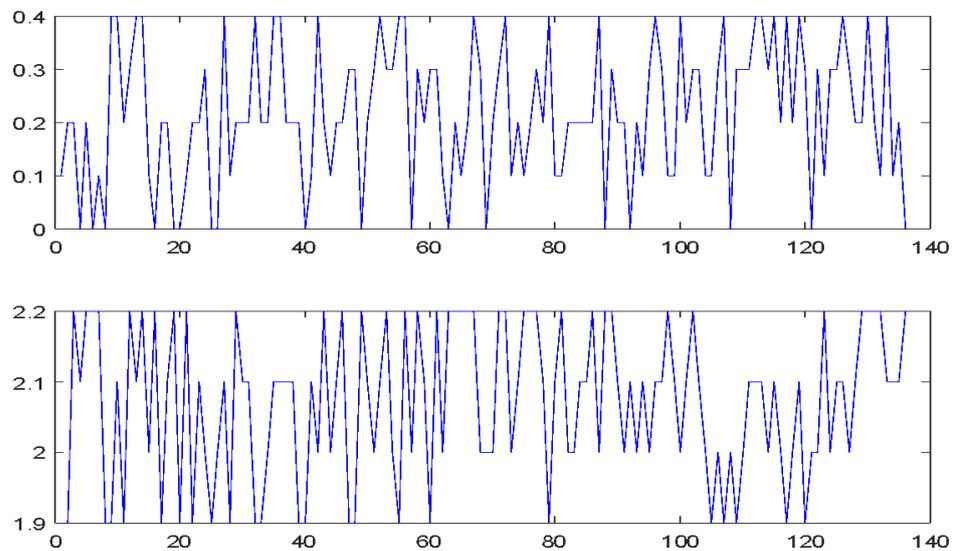


Рис. 8. Пример выборки процентной загрузки ресурсов (сверху вниз соответственно – сеть, процессор) от дискретов времени (временные отчеты от 0 до 140) для состояния Z_1

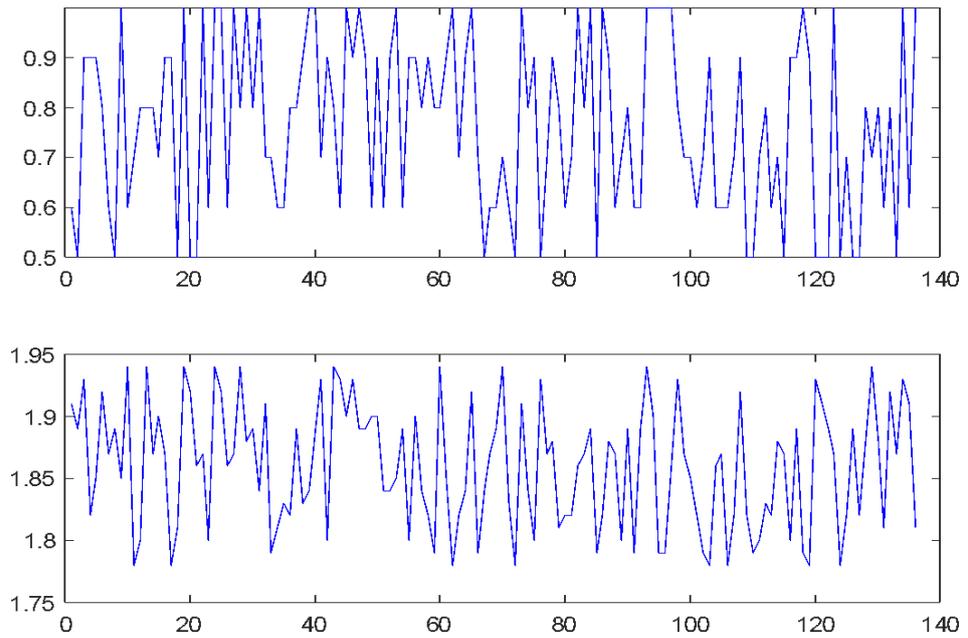


Рис. 9. Пример выборки процентной загрузки ресурсов (сверху вниз соответственно – сеть, процессор) от дискретов времени (временные отчеты от 0 до 140) для состояния Z_2

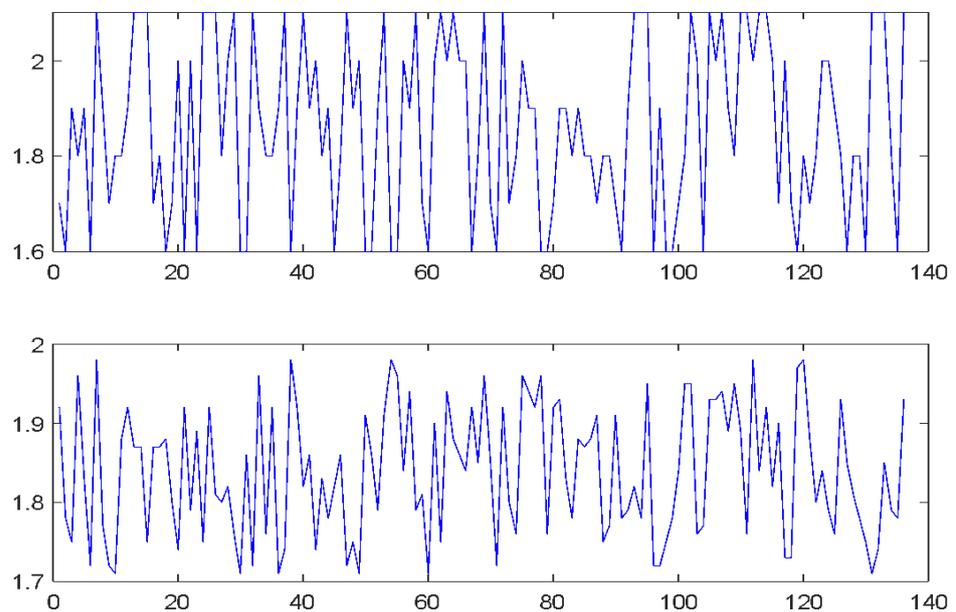


Рис. 10. Пример выборки процентной загрузки ресурсов (сверху вниз соответственно – сеть, процессор) от дискретов времени (временные отчеты от 0 до 140) для состояния Z_3

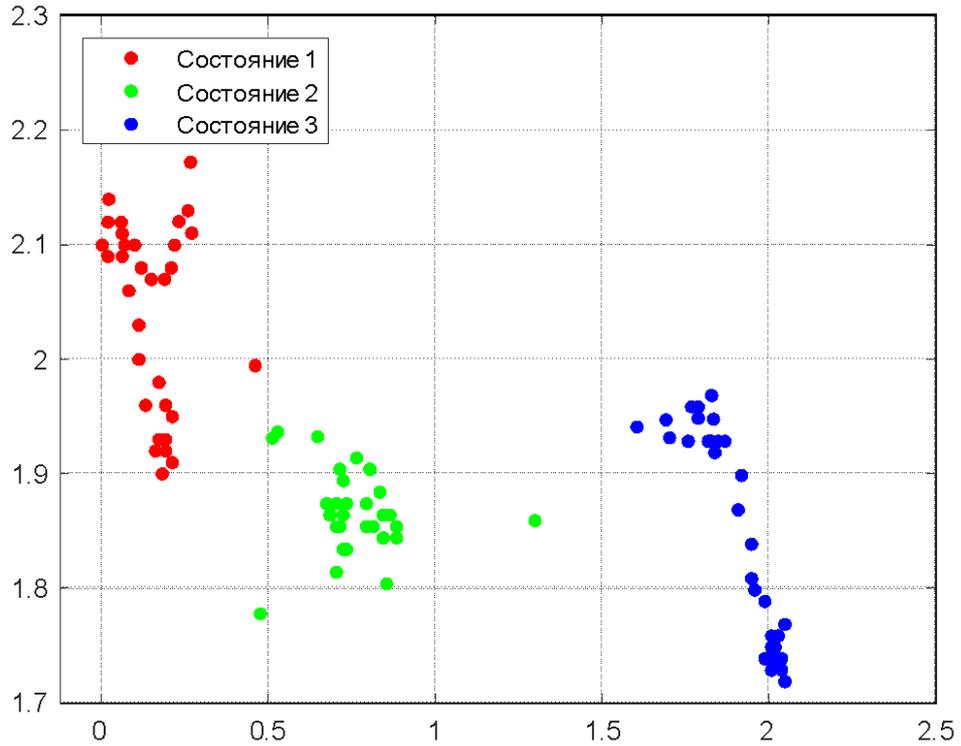


Рис. 11. Результаты состояний на двумерных координатных осях

В ходе эксперимента рассматривались алгоритмы классификации a_j входного вектора X_i для отображения $Z \rightarrow C$. Независимо друг от друга обученные и работающие классификаторы $\{a_1, a_2, \dots, a_k\} \in a$ $k = 4$ (наивный байесовский классификатор, деревья решений, дискриминантный анализ, метод k ближайших соседей) выдавали последовательности результатов

$$Z = \{(z_{a_1}^{c_0}, z_{a_1}^{c_1}, \dots, z_{a_1}^{c_n}), (z_{a_2}^{c_0}, z_{a_2}^{c_1}, \dots, z_{a_2}^{c_n}), \dots, (z_{a_j}^{c_0}, z_{a_j}^{c_1}, \dots, z_{a_j}^{c_n}), \dots, (z_{a_k}^{c_0}, z_{a_k}^{c_1}, \dots, z_{a_k}^{c_n})\}.$$

Результирующий класс c_i состояния z_i , предсказываемый каждой моделью, определяется усреднением значений вычисленных вероятностей:

$$a_{c_i} = \frac{1}{K} \sum_{k=1}^K w_k a_k(x_i)$$

Применим ряд «слабых» заранее обученных на размеченной выборке классификаторов a_R : наивный байесовский классификатор, деревья решений, дискриминантный анализ, метод k ближайших соседей. Области их оценки приведены на рисунке 12.

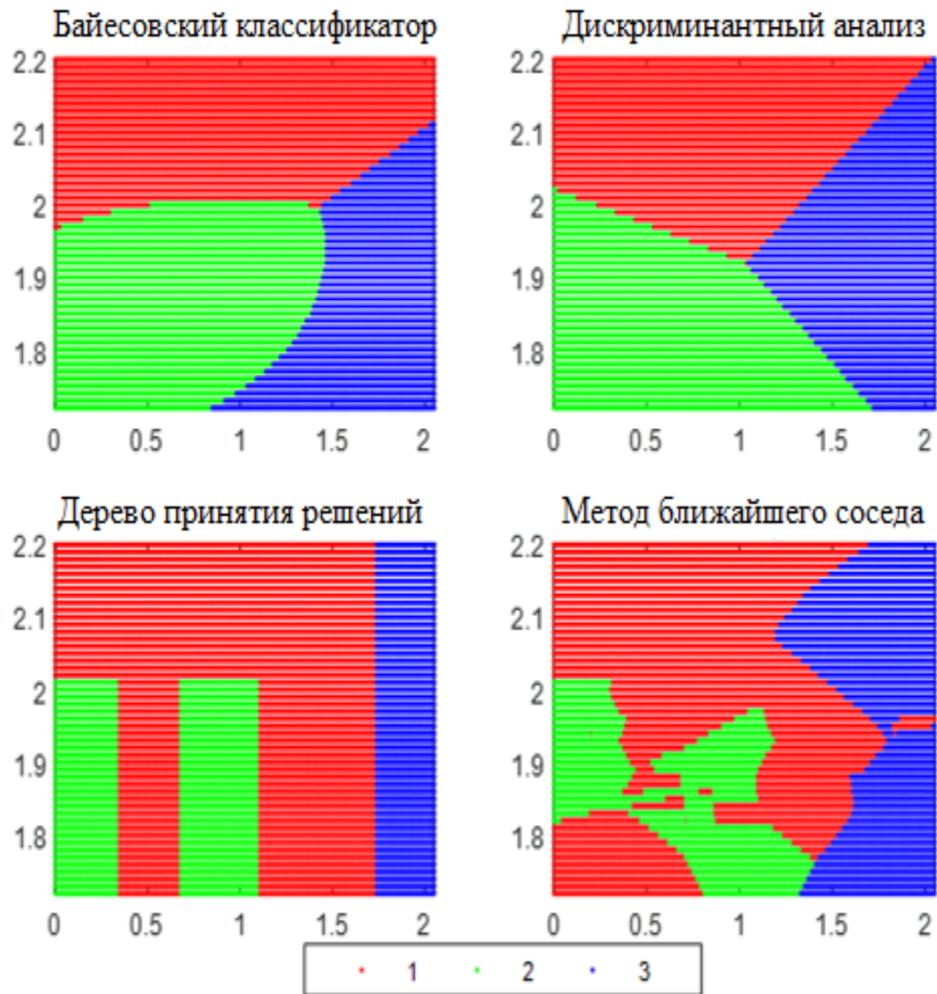


Рис. 12. Области оценки классификаторов a_{R_i}

Получив эти данные, строим таблицу по ним. Получим вероятности ошибочной классификации, полученные в результате применения «слабых» классификаторов a_{C_i}

В табл. 1 представлены вероятности ошибочной классификации, полученные в результате применения «слабых» классификаторов a_{C_i} .

Таблица 1.

Вероятность ошибочно классифицированных значений выборки

	Кластер 1	Кластер 2	Кластер 3	Всего для выборки
Наивный байесовский классификатор	0,18	2	2	0,07
Дискриминантный анализ	0,16	2	2	0,07
Дерево принятия решений	0,08	4	4	0,05
Метод ближайшего соседа	0,2	8	6	0,11

На рис. 13 представлена визуализация вероятностной оценки ошибочной классификации.

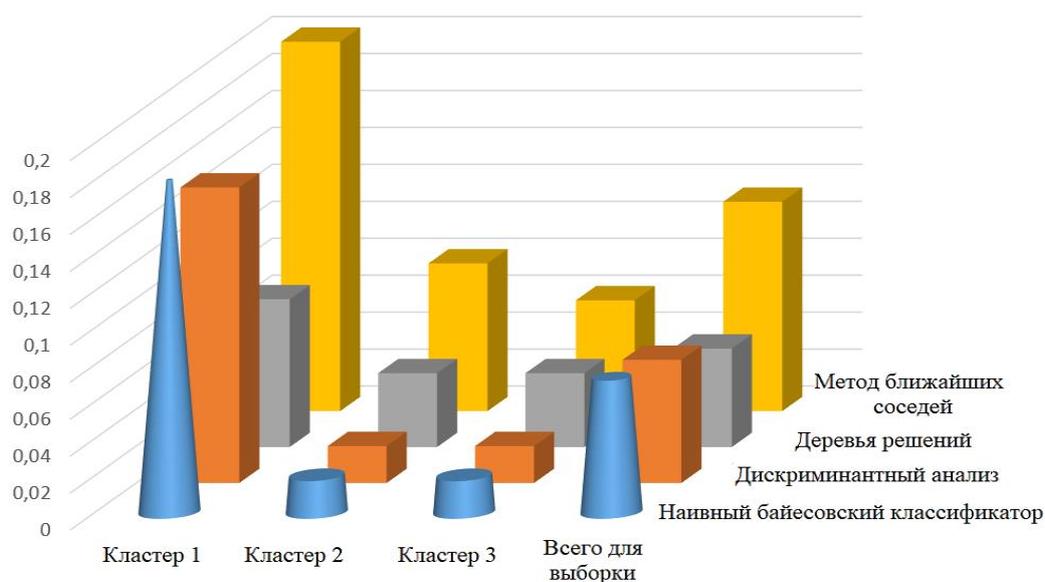


Рис. 13. Вероятности ошибочной классификации

Полученный подход на основе экспериментальных данных дает общую точность классификации состояний на уровне 0,93. Следует отметить, что данные не были предварительно обработаны, не были очищены от шумов, а также были получены с низкой частотой дискретизации.

Таким образом, данное решение позволяет классифицировать текущее состояние и может быть использовано в качестве теоретической базы для интеграции методов машинного обучения в анализе состояния устройств бортовых систем.

ЗАКЛЮЧЕНИЕ

Данная выпускная квалификационная работа была посвящена идентификации состояний устройств бортовых систем с использованием метода бэггинга.

В процессе исследования были проведены следующие этапы:

1. Изучение бортовых систем и их устройств, а также их классификация и сферы использования этих систем.
2. Была дана оценка состоянию функционирования устройств бортовых систем.
3. Изучение существующих подходов и методов идентификации состояний устройств бортовых систем на водном судне.
4. Анализ и оценка полученных результатов и сравнение с результатами, полученными на основе других методов машинного обучения.

В результате исследования были получены следующие результаты:

1. Был разработан метод идентификации состояний устройств бортовых систем на основе метода бэггинга.
2. Были разработаны и протестированы алгоритмы машинного обучения с применением метода бэггинга, позволяющие достигать высокой точности идентификации состояний устройств бортовых систем.
3. Был проведен анализ и оценка полученных результатов и сравнение с результатами, полученными на основе других методов машинного обучения.
4. Полученные результаты показали, что использование метода бэггинга позволяет повысить точность идентификации состояний устройств бортовых систем.

Таким образом, результаты исследования показали, что применение метода бэггинга может быть эффективным подходом для идентификации состояний устройств бортовых систем. Разработанные алгоритмы машинного обучения с применением метода бэггинга могут быть использованы для

создания более сложных систем мониторинга и контроля состояния бортовых систем.

Дальнейшие исследования могут быть направлены на улучшение качества методов идентификации состояний устройств бортовых систем на водном судне с использованием метода бэггинга. Также можно рассмотреть возможность применения других методов машинного обучения в сочетании с методом бэггинга для улучшения точности идентификации состояний устройств бортовых систем на водном судне.

В целом, результаты данной выпускной квалификационной работы могут быть полезны для разработки новых систем мониторинга и контроля состояния бортовых систем, что в свою очередь, способствует повышению безопасности плавания и экономической эффективности работы судна.

Источники

1. Сухопаров, М.Е. Анализ состояния устройств интернета вещей на основе бэггинга / М.Е. Сухопаров, И.С. Лебедев // Т-Comm: Телекоммуникации и Транспорт. – 2020. – Т. 14. – №12. – С. 45–50.
2. Сухопаров, М.Е. Применение ансамбля обученных на несбалансированных выборках нейросетей при анализе состояния устройств интернета вещей / М.Е. Сухопаров, И.С. Лебедев // Проблемы информационной безопасности. Компьютерные системы. – 2021. – № 2 (46). – С. 127–134.
3. Сухопаров М. Е., Семенов В. В., Салахутдинова К. И., Лебедев И. С. Выявление аномального функционирования устройств «Индустрии 4.0» на основе поведенческих паттернов // Проблемы информационной безопасности. Компьютерные системы. 2020. № 1 (41). С. 96-102.
4. Зикратов И. А., Зикратова Т. В., Лебедев И. С. Доверительная модель информационной безопасности мультиагентных робототехнических систем с децентрализованным управлением // Научно-технический вестник информационных технологий, механики и оптики. 2014. № 2 (90). С. 47-52.
5. Gao D., Reiter M., Song D. Beyond output voting: Detecting compromised replicas using HMM-based behavioral distance. IEEE Transactions on Dependable and Secure Computing, 2009. Vol. 6, no. 2. P. 96-110.
6. Devesh M., Kant A. K., Suchit Y. R., Tanuja P., Kumar S. N. Fruition of cps and iot in context of Industry 4.0. Intelligent Communication, Control and Devices. Advances in Intelligent Systems and Computing, 2020. Vol. 989. P. 367-375.
7. Семенов В. В., Лебедев И. С., Сухопаров М. Е. Подход к классификации состояния информационной безопасности элементов киберфизических систем с использованием побочного электромагнитного излучения //

- Научно-технический вестник информационных технологий, механики и оптики. 2018. № 1. С. 98-105.
8. Сошникова Л. А., Тамашевич В. Н., Усбе Г., Шефер М. Многомерный статистический анализ в экономике: учебное пособие для вузов. М.: ЮНИТИ – Дана, 1999. 598 с.
 9. НД N 2-020101-114 Правила классификации и постройки морских судов. Часть XI. Электрическое оборудование. 2019. С. 78-116.
 10. В.И.Меньшиков, В.В.Ковальчук, А.Н.Папуша, И.А.Кулезнёв. Оценка эффективности процедур по защите безопасности судна. Морская академия МГТУ, кафедра судовождения Институт экономики, управления и международных отношений МГТУ, кафедра информационных систем и прикладной математики, Политехнический институт МГТУ, кафедра механики сплошных сред и морского нефтегазового дела. 2013.
 11. ГОСТ Р 56243-2014. Внутренний водный транспорт. Требования по обеспечению надежности механизмов [Электронный ресурс]. - Электронный фонд правовых и нормативно-технических документов. – 2015. – Режим доступа: <https://docs.cntd.ru/document/1200114751>.
 12. ГОСТ 32455-2013. Глобальная навигационная спутниковая система. Общие требования, методы и требуемые результаты испытаний [Электронный ресурс]. - Электронный фонд правовых и нормативно-технических документов. – 2014. – Режим доступа: <https://docs.cntd.ru/document/1200107398>.
 13. ГОСТ Р ИСО 10006–2019 “Менеджмент качества. Руководящие указания по менеджменту качества в проектах” [Электронный ресурс]. - Электронный фонд правовых и нормативно-технических документов. – 2014. – Режим доступа: <https://docs.cntd.ru/document/1200167119>.
 14. К.В.Воронцов. Математические методы обучения по прецедентам (теория обучения машин) С. 18-39.

15. К.В.Воронцов. Математические методы обучения по прецедентам (теория обучения машин) С. 42-48.
16. К.В.Воронцов. Математические методы обучения по прецедентам (теория обучения машин) С. 51-78
17. К.В.Воронцов. Математические методы обучения по прецедентам (теория обучения машин) С. 80-99
18. Ю.С.Кашницкий, Д. И. Игнатов. Ансамблевый метод машинного обучения, основанный на рекомендации классификаторов. С. 37-41.
19. ГОСТ 27.310–95. Надежность в технике. Анализ видов, последствий и критичности отказов. Основные положения [Электронный ресурс]. – Электронный фонд правовых и нормативно-технических документов. – 1995. – Режим доступа: <https://docs.cntd.ru/document/1200001363>.
20. ГОСТ 27.002–2015. Надежность в технике. Термины и определения [Электронный ресурс]. – Электронный фонд правовых и нормативно-технических документов. – 2015. – Режим доступа: <https://docs.cntd.ru/document/1200136419>.
21. ГОСТ Р 27.003–2011. Надежность в технике. Управление надежностью. Руководство по заданию технических требований к надежности [Электронный ресурс]. – Электронный фонд правовых и нормативно-технических документов. – 2011. – Режим доступа: <https://docs.cntd.ru/document/1200095898>.