



НАУКА И ПРАКТИКА: ПРОЕКТНАЯ ДЕЯТЕЛЬНОСТЬ — ОТ ИДЕИ ДО ВНЕДРЕНИЯ

Материалы XI региональной
научно-практической конференции
Томск, 2022

Министерство науки и высшего образования Российской Федерации

Томский государственный университет
систем управления и радиоэлектроники

**НАУКА И ПРАКТИКА:
ПРОЕКТНАЯ ДЕЯТЕЛЬНОСТЬ –
ОТ ИДЕИ ДО ВНЕДРЕНИЯ**

**Материалы XI региональной
научно-практической конференции
Томск, 2022**

Томск
Издательство ТУСУРа
2022

УДК 336.114(063):005.8
ББК 94.3
Н34

Организационный комитет конференции

Сенченко П.В. – канд. техн. наук, доцент, проректор по учебной работе
Лощилов А.Г. – канд. техн. наук, доцент, проректор по научной работе
и инновациям

Богомоллова А.В. – канд. экон. наук, декан ЭФ

Михальченко С.Г. – д-р техн. наук, доцент, зав. кафедрой ПрЭ

Незнамова Е.Г. – канд. биол. наук, доцент, доцент каф. РЭТЭМ

Пахмурин Д.О. – канд. техн. наук, доцент, доцент кафедры ПрЭ

Покровская Е.М. – канд. филос. наук, доцент, зав. кафедрой ИЯ

Раитина М.Р. – д-р филос. наук, доцент, доцент кафедры ФиС

Рахманенко И.А. – канд. техн. наук, доцент, доцент кафедры КИБЭВС

Каранский В.В. – канд. техн. наук, декан ФЭ

Сидоров А.А. – канд. экон. наук, зав. кафедрой АОИ

Убайчин А.В. – канд. техн. наук, доцент, доцент кафедры РСС

Хатьков Н.Д. – канд. техн. наук, доцент, доцент кафедры СВЧиКР

Н34 **Наука и практика: проектная деятельность – от идеи до внедрения:**
материалы X региональной науч.-прак. конф., Томск, 2022. – Томск :
Томск. гос. ун-т систем упр. и радиоэлектроники, 2022. – 802 с.

Представлены результаты реализации проектов школьников, студентов и руководителей научно-исследовательской работы учащихся, в рамках проектных групп или индивидуальных научных исследований, имеющих инновационную составляющую и ориентированных на дальнейшее коммерческое использование. Основной целью интернет-конференции является обмен информацией о новых научных направлениях, инновационных подходах и методах решения актуальных проблем, а также представление и обсуждение результатов исследований.

Содержатся доклады, связанные с радиоэлектроникой, радиотехникой, нанотехнологиями, приборостроением, энергетикой и силовой электроникой, радиосвязью и СВЧ, автоматизированными системами обработки информации, а также биомедицинскими, экономическими, социальными и информационными технологиями.

УДК 336.114(063):005.8
ББК 94.3

© Томск. гос. ун-т систем упр.
и радиоэлектроники, 2022

3. Пять перспективных языков программирования со светлым будущим [Электронный ресурс]. URL: <https://habr.com/ru/post/310252> (дата обращения: 03.11.2022).

ПРОВЕРКА БЛАГОНАДЁЖНОСТИ ИНДИВИДУАЛЬНЫХ ПРЕДПРИНИМАТЕЛЕЙ С ИСПОЛЬЗОВАНИЕМ МЕТОДА СЛУЧАЙНОГО ЛЕСА

Е.И. Васильев, А.В. Осипенко, Г.Р. Егле, студенты каф. БИС

г. Томск, ТУСУР, egg.or.no@gmail.com

Научный руководитель: А.С. Колтайс, ст. преподаватель каф. КИБЭВС

г. Томск, ТУСУР, kas@fb.tusur.ru

Проект ГПО КИБЭВС-1703 – Методика работы с системой СПАРК

Описывается применение метода машинного обучения случайный лес для проверки благонадёжности индивидуальных предпринимателей. Для него проводится подбор наиболее подходящих параметров и сравнительный анализ по важным критериям машинного обучения, влияющим на будущую автоматизированную систему. Результаты исследования будут способствовать развитию методов и систем по определению благонадёжных контрагентов.

Ключевые слова: *машинное обучение, автоматизация, благонадёжность, контрагент*

На сегодняшний день в современных условиях рыночной конкуренции, где каждый стремится получить коммерческую выгоду и развить своё предприятие, актуальной является проблема поиска и определения благонадёжного контрагента. В 2021 году по данным федеральной службы информации количество индивидуальных предпринимателей составило 3621580 человек, а количество юридических лиц 3316168 [1]. Так в 2022 году количество индивидуальных предпринимателей выросло в среднем на 8,2 % по сравнению с предыдущим годом. Все представленные данные говорят о том, что с каждым годом появляется всё больше ИП и ЮЛ, а следовательно, становится сложнее исключить неблагонадёжных и найти себе наиболее подходящих контрагентов.

Таким образом, появляется острая необходимость в надёжном и специализированном продукте, который позволит быстро, а главное точно проверять контрагентов. Для создания такой системы необходимо определить точные методы и модели проверки индивидуальных предпринимателей, которые бы смогли провести корректную обработку больших массивов данных.

Одним из вариантов достижения поставленной цели является использование машинного обучения, которое объективно оценивает риски работы с тем или

иным контрагентом, что позволяет быстро и точно обрабатывать информацию в автоматизированном формате.

Для оценки благонадёжности ИП уже использовались методы машинного обучения, среди которых выделяется «Дерево принятия решений» [2]. Он проявляет себя как по основополагающему критерию – точности, так и по вторичным: скорости анализа и весу обученной модели. На основе данных результатов существует возможность улучшить модель ИП – использовать ансамблевый метод «Случайный лес», который состоит из нескольких деревьев принятия решений [3].

Данный метод проводит деление тренировочного набора данных на несколько частей, с помощью которых происходит обучение отдельных деревьев. При анализе каждое дерево может предоставить один из результатов, который был получен посредством последовательного ответа на вопросы вида «Да/Нет». Стоит отметить, что метод дерева принятия решений подготавливает разветвления под выборку, из-за чего возрастает риск переобучения и как итог плохого результата анализа. Для решения этой проблемы необходимо подбирать параметры глубины, то есть количество разветвлений до получения результата, или использовать «Случайный лес».

Исследуемый метод имеет в названии «случайный» как раз за счёт создания деревьев с разными параметрами и как итог избегает переобучения модели. При использовании нескольких одинаковых «деревьев», обученных на разных частях выборки, будет просто «лес» без введения хаотичности в параметрах.

Для анализа данного метода машинного обучения использовалась выборка из 7653 индивидуальных предпринимателей Томской области. Входными данными послужили критерии, отобранные в соответствии с нормативно-правовыми актами РФ [4].

Несмотря на случайный подбор параметров деревьев внутри леса, следует обратить внимание на устанавливаемые ограничения диапазона значений. Так определяются гиперпараметры, такие как максимальная глубина дерева, минимальное количество элементов для разделения узла, используется ли весь набор данных при обучении деревьев, критерий для оценки качества разделения узлов (загрязнение Джини или энтропия).

Для примера, в таблице 1 приведена средняя оценка максимальной глубины деревьев по метрикам на долю правильных ответов (ассигасу), точность (precision), полноту (recall) и на среднюю гармоническую точности и полноты (F1). У показателя параметра глубины до 5 пунктов имеет ассигасу ниже 99%, что хуже результатов «Дерева принятия решений».

Несмотря на лучшую в рамках таблицы метрику F1 у параметров 8 и 10, при комбинации с другими настройками максимальная глубина в 9 пункте будет с лидирующим значением. Из других подходящих параметров можно выделить

количество деревьев в 100 объектов, использование загрязнения Джини в качестве оценки качества разделения по сравнению с энтропией, а также слабое влияние гиперпараметра, определяющего обучение деревьев на всём наборе данных или на разных его частях.

Таблица 1 – Сравнение результатов по максимальной глубине деревьев

Максимальная глубина	Accuracy	Precision	Recall	F1
3	97,143	97,113	99,956	98,513
4	98,896	98,854	99,967	99,453
5	99,924	99,912	99,967	99,946
6	99,928	99,948	99,969	99,961
7	99,929	99,956	99,969	99,961
8	99,926	99,957	99,969	99,963
9	99,932	99,955	99,969	99,960
10	99,928	99,957	99,968	99,963
11	99,932	99,955	99,969	99,962
12	99,928	99,955	99,971	99,961

Итоговые показатели модели были получены со значениями исследованных настроек и с использованием кросс-валидации. Точность модели была повышена по сравнению с «Деревом принятия решений», так как с новым методом было допущено только две ошибки из всей выборки, но в то же время, по остальным критериям сравнения он уступает прошлому (таблица 2).

Таблица 2 – Сравнение параметров двух методов

Критерий оценивания	Дерево принятия решений	Случайный лес
Время обучения, мкс	3 042	74 403
Время анализа, мкс	552	13 596
Доля правильных ответов, %	99,93	99,974
Вес модели, КБ	3,1	188,2
Визуализация	возможна	возможна частично

Время обучения является маловажным критерием для автоматизированной системы, но время анализа – наоборот, будет влиять на работу сервера во время ответа. Этот параметр стоит учитывать уже в соответствии со всей работоспособностью системы, например, при расчёте производительности.

Вес модели хоть и был увеличен более чем в 60 раз, но всё ещё является довольно незначительным критерием в рамках современных технологий хранения. Визуализация для случайного леса становится под вопросом, так как

отобразить все 100 деревьев возможно, но для пользователя это будет не самым лучшим опытом.

Исходя из полученных результатов, новая модель по проверке благонадёжности ИП смогла улучшить результаты по точности при помощи использования «Случайного леса» (99,974%) по сравнению с «Деревом принятия решений» (99,93%). Внедрение данного метода в автоматизированную систему будет зависеть и от других выявленных показателей, но в первую очередь он должен быть надёжным, что и было продемонстрировано в результатах данной работы.

Литература

1. Федеральная служба информации РФ – Статистическая информация. URL: <https://xn--h1ari.xn--plai/Main/StatisticalInformation> (дата обращения: 26.10.2022).
2. Гриценко А.А., Васильев Е.И., Мануилова Д.Е. Сравнительный анализ методов машинного обучения для определения благонадёжности индивидуальных предпринимателей // XIX междунар. конф. студентов, аспирантов и молодых ученых «Перспективы развития фундаментальных наук», Томск, 26–29 апреля 2022 г. Томск, 2022. С. 17–19.
3. Официальная документация sklearn по случайному лесу. URL: <https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html> (дата обращения: 26.10.2022).
4. Козлова Н.А., Колтайс А.С., Устинов А.О. Модель оценки благонадежности индивидуальных предпринимателей // Сб. науч. тр. XVIII междунар. конф. студентов, аспирантов и молодых ученых «Перспективы развития фундаментальных наук». Томск, 2021. С. 53–55.

ТЕХНОЛОГИЯ ВЫДЕЛЕНИЯ ЖЕЛЕЗНОДОРОЖНЫХ ПУТЕЙ (РЕЛЬС) ПО ИЗОБРАЖЕНИЯМ, ПОЛУЧЕННЫМ С ПОМОЩЬЮ БПЛА

Г.А. Игнатеня, К.С. Емельянов, студенты каф. АСУ

г. Томск, ТУСУР, gennadyignatenya@yandex.ru

Научный руководитель: М.Ю. Катаев, д-р техн. наук, проф. каф. АСУ ТУСУР

*Проект ГПО АСУ-2201 – Мониторинг территории
по изображениям БПЛА*

Приводится описание технологии выделения рельс по изображениям, полученным с помощью БПЛА. Для разработки технологии применяются подходы компьютерного зрения и нейронных сетей. Приведена обработка реальных изображений и выполнен анализ результатов.

Содержание

Секция 1. АЛГОРИТМЫ И АВТОМАТИЗИРОВАННЫЕ СИСТЕМЫ ОБРАБОТКИ ИНФОРМАЦИИ И УПРАВЛЕНИЯ

А.Р. Мальцев, К.А. Яковлев, Д.А. Агафонов АНАЛИЗ И ОБРАБОТКА СПУТНИКОВЫХ ИЗОБРАЖЕНИЙ SENTINEL-2 ДЛЯ СЕЛЬСКОГО ХОЗЯЙСТВА.....	3
И.Ю. Атрошенко, Н.С. Альшевский АЛГОРИТМ ДЛЯ ЛЕКСИКОГРАФИЧЕСКОГО РАЗБОРА ФАЙЛА СОГЛАСНО ЗАДАННОМУ ШАБЛОНУ	5
М.А. Беляева ИНФОРМАЦИОННАЯ СИСТЕМА ВИЗУАЛИЗАЦИИ СПУТНИКОВЫХ ДАННЫХ О СОДЕРЖАНИИ ПАРНИКОВЫХ ГАЗОВ. БАЗА ДАННЫХ	8
В.С. Чебанов ИССЛЕДОВАНИЕ ВЛИЯНИЯ КОЛИЧЕСТВА ЭПОХ ТРАНСФЕРНОГО ОБУЧЕНИЯ НА ВЕРОЯТНОСТЬ РАСПОЗНАНИЯ С ИСПОЛЬЗОВАНИЕМ МИКРОКОМПЬЮТЕРА JETSON NANO	11
К.В. Васильев, Д.О. Черникова, А.В. Курдачева ОТ ИГРЫ ДО МАКЕТА УМНОГО ДОМА	14
Б.Б. Джумабаев ИНСТРУМЕНТАРИЙ РАЗРАБОТКИ ВЕБ-ПРИЛОЖЕНИЯ РАЗВИТИЯ УНИВЕРСАЛЬНЫХ КОМПЕТЕНЦИЙ СТУДЕНТОВ	17
Е.И. Васильев, А.В. Осипенко, Г.Р. Егле ПРОВЕРКА БЛАГОНАДЁЖНОСТИ ИНДИВИДУАЛЬНЫХ ПРЕДПРИНИМАТЕЛЕЙ С ИСПОЛЬЗОВАНИЕМ МЕТОДА СЛУЧАЙНОГО ЛЕСА	21
Г.А. Игнатеня, К.С. Емельянов ТЕХНОЛОГИЯ ВЫДЕЛЕНИЯ ЖЕЛЕЗНОДОРОЖНЫХ ПУТЕЙ (РЕЛЬС) ПО ИЗОБРАЖЕНИЯМ, ПОЛУЧЕННЫМ С ПОМОЩЬЮ БПЛА.....	24
Д.Р. Клецко, Г.П. Ермак РАЗРАБОТКА ГОЛОГРАФИЧЕСКОГО ДИСПЛЕЯ.....	28
А.Б. Кураков, А.Б. Гомбоин РАЗРАБОТКА КРОССПЛАТФОРМЕННОГО ПРИЛОЖЕНИЯ ДЛЯ НАХОЖДЕНИЯ ЦЕНТРА ТЯЖЕСТИ ЧЕЛОВЕКА.....	31
А.Э. Скиба, Е.В. Истомин МОДЕЛЬ КЛИЕНТ-СЕРВЕРНОЙ АРХИТЕКТУРЫ ДЛЯ ИНФОРМАЦИОННОЙ СИСТЕМЕ ПО МОНИТОРИНГУ СОСТОЯНИЯ ЗДОРОВЬЯ СОТРУДНИКОВ	34
А.С. Гаврилов, В.А. Кикоть, Р.М. Муромцев СЕРВИС ДЛЯ ИДЕНТИФИКАЦИИ КЛЮЧЕВЫХ ПОЛЬЗОВАТЕЛЕЙ В СОЦИАЛЬНОЙ СЕТИ.....	36