

МЕЖДИСЦИПЛИНАРНЫЕ ИССЛЕДОВАНИЯ

А.П. Казун

*к.с.н., директор Института анализа предприятий и рынков,
Национальный исследовательский университет «Высшая школа
экономики» (Москва)*

В.Ю. Девятников

*стажёр-исследователь, Национальный исследовательский университет
«Высшая школа экономики» (Москва)*

М.Д. Белов

аспирант, Калифорнийский университет (Лос-Анджелес, США)

АСИММЕТРИЯ ИНФОРМАЦИИ НА РЫНКЕ АДВОКАТСКИХ УСЛУГ: СОЗДАНИЕ РЕЙТИНГА МЕТОДАМИ МАШИННОГО ОБУЧЕНИЯ¹

Аннотация. Как понять, хорош адвокат или нет? В данной статье мы описываем методологию, позволяющую измерять качество юридической помощи через разницу между тем, что предсказывает модель машинного обучения по объективным характеристикам дела, и тем, чем дело закончилось в реальности. Мы исходим из предпосылки, что если у адвоката дела стабильно заканчиваются лучше ожидаемого, то он делает свою работу хорошо. Похожую логику давно используют в других областях: так оценивают учителей в экономике образования, хирургов в медицине и игроков в профессиональном спорте. Мы адаптируем эти подходы для уголовного судопроизводства. Такая прогнозная модель может быть обучена на данных российских судов за 2010–2025 гг. и будет учитывать, что уголовное дело проходит через несколько последовательных процессуальных развилок. Рейтинг корректируется на командную работу нескольких защитников, стадию рассмотрения и малое число дел. В статье мы обсуждаем слабые места подхода: прежде всего то, что досудебная стадия для нас невидима, публикуемые решения неполны, а сильные адвокаты могут целенаправленно браться за сложные дела. Тем не менее даже в таком виде подобный инструмент оценки способен дать клиентам значимый и объективный ориентир при выборе защитника, что может существенно улучшить качество информации на рынке юридических услуг.

Ключевые слова: оценка эффективности работы адвокатов, прогнозирование судебных решений, машинное обучение, информационная асимметрия, уголовная защита.

JEL: K14, K41, C53, C55, D82

УДК: 34.08, 34.096

DOI: 10.52342/2587-7666VTE_2026_3_115_135

© А.П. Казун, В.Ю. Девятников, М.Д. Белов, 2026

© ФГБУН Институт экономики РАН «Вопросы теоретической экономики», 2026

ДЛЯ ЦИТИРОВАНИЯ: Казун А.П., Девятников В.Ю., Белов М.Д. Асимметрия информации на рынке адвокатских услуг: создание рейтинга методами машинного обучения // Вопросы теоретической экономики. 2026. №3. С. 115–135. DOI: 10.52342/2587-7666VTE_2026_3_115_135.

FOR CITATION: Kazun A., Devyatnikov V., Belov M. Information Asymmetry in the Legal Services Market: Creating a Ranking Using Machine Learning Methods // Voprosy teoreticheskoy ekonomiki. 2026. No. 3. Pp. 115–135. DOI: 10.52342/2587-7666VTE_2026_3_115_135.

¹ Исследование выполнено в рамках Программы фундаментальных исследований НИУ ВШЭ, № HSE-BR-2025-034.

Введение

Оценка качества работы адвокатов в уголовном судопроизводстве — одна из наиболее сложных нерешённых задач в экономике права. В отличие от многих других профессий, где результативность специалиста может быть измерена с помощью объективных показателей — количество успешных операций у хирурга, результаты учеников у преподавателя, статистика голов или результативных передач у спортсмена, — работа адвоката оценивается преимущественно на основе субъективных критериев: репутации, отзывов, стоимости услуг. Между тем рынок юридических услуг в сфере уголовной защиты характеризуется выраженной информационной асимметрией, при которой потенциальный доверитель не обладает достаточными инструментами для объективного сопоставления профессиональных качеств различных адвокатов [Galanter, 1974].

Проблема информационной асимметрии на рынке юридических услуг приобретает особую остроту в контексте уголовного правосудия, где последствия выбора защитника для подсудимого наиболее значимы. Ошибка в выборе адвоката может привести к реальному лишению свободы в тех случаях, когда более квалифицированный защитник мог бы добиться или условного наказания, или прекращения дела, или оправдания. Зарубежные исследования показывают, что разброс в результативности адвокатов существен: лучшие из них снижают вероятность лишения свободы на 14 п. п. по сравнению с адвокатами, котирующимися на уровне нижнего, десятого, процентиля [Abrams, Yoon, 2007]. Однако для клиента, обращающегося за помощью на рынок юридических услуг, эта информация остаётся недоступной.

Существующие подходы к оценке качества юридической помощи в уголовных делах можно разделить на несколько групп. Первая — простые показатели процентов выигрышных дел (win rates), которые не учитывают различия в сложности дел и потому систематически искажают картину. Вторая — экспертные оценки коллег и судей, которые, будучи информативными, страдают от субъективности и ограниченного масштабирования. Третья — рейтинговые системы, основанные на репутации юридических фирм, корреляция которых с реальными результатами, если верить последним исследованиям, оказывается крайне слабой [Mojon, Mahari, Lera, 2025]. Ни один из этих подходов не предлагает систематической оценки, основанной на методологии, которая бы контролировала сложность дел, т.е. обеспечивала бы сопоставление «при прочих равных условиях».

Методологические трудности оценки качества адвокатов во многом аналогичны проблемам, возникающим при оценке качества работы специалистов в других областях. В сфере образования исследователи используют модели добавленной стоимости учителя. Одна из проблем этой сферы состоит в том, что успех ученика зависит совсем не только от учителя. Нельзя не учитывать и множество других характеристик — от базовой успеваемости и мотивации к обучению, способностей ученика, образования его родителей и т.п. Но есть модели, позволяющие оценить вклад конкретного преподавателя в достижения учеников, контролируя исходные характеристики последних [Chetty, Friedman, Rockoff, 2014a; Chetty, Friedman, Rockoff, 2014b]. Аналогичные модели учитывающие то, что успешность операции зависит не только от хирурга, но и от того, насколько в целом здоровым был человек, который ложится на операционный стол, используются в медицине [Hall et al., 2015; Krell et al., 2014]. Наконец, в спортивной аналитике метрика, называемая «Adjusted Plus-Minus» (APM), позволяет оценить индивидуальный вклад игрока в командный результат, контролируя силу партнёров и соперников [Grassetti et al., 2021; Macdonald, 2012]. Общий принцип всех этих подходов состоит в сравнении фактических результатов с ожидаемыми (предсказанными) на основе характеристик контекста, в котором действует профессионал: разница между фактическим и предсказанным исходом атрибутируется как показатель качества работы специалиста.

В данной статье мы предлагаем методологию, которая использует описанный выше подход, применяющийся при оценке специалистов в других сферах деятельности. Суть состоит в том, что мы строим модели, предсказывающие судебные решения на основе различных характеристик дела, на которые адвокат не может повлиять, а затем сравниваем их с реальными исходами. Разница между реальностью и прогнозом будет рассматриваться как вклад адвоката: если он систематически добивается более благоприятных исходов, то мы увеличиваем его рейтинг. Разумеется, это справедливо при условии, что для каждого адвоката мы оцениваем большое количество дел за достаточно длительный период, ибо на «длинной дистанции» некоторые случайные факторы (такие как «удачное стечение дел») сглаживаются, а накопленная дельта начинает отражать реальный уровень эффективности защитника.

Применительно к российскому контексту предлагаемая методология приобретает особую значимость в силу нескольких обстоятельств. Во-первых, российское уголовное правосудие характеризуется так называемым обвинительным уклоном: доля оправдательных приговоров составляет менее 0,2% [Панеях, 2012]. Этот факт делает некорректным простое сравнение процента оправдательных приговоров между адвокатами (ввиду редкости такого исхода). При столь низком базовом уровне оправданий необходимо оценивать более тонкие градации исходов. Во-вторых, тексты судебных решений размещаются в открытом доступе в соответствии с Федеральным законом № 262-ФЗ от 2008 г., что обеспечивает масштабную информационную базу для количественного анализа. Суммарный объём доступных решений за период 2010–2025 гг. превышает 7 млн документов. В-третьих, такие факторы, как присутствие прокурора в процессе, существенно влияют на исходы дел [Volkov, 2021]. Это делает необходимым тщательный контроль разнообразных характеристик дела при любой попытке приписать благоприятные исходы именно действиям защитника.

Вклад настоящей статьи в анализ рассматриваемой проблемы можно охарактеризовать как двойной. С теоретической точки зрения статья предлагает формализованную методологию оценки эффективности работы адвокатов, основанную на подходе *value-added*, который до настоящего времени систематически не применялся в данном контексте. Методология включает трёхэтапную модель прогнозирования судебных решений, систему корректировок для командной работы и малых выборок, а также байесовское сглаживание рейтинговых оценок. С практической точки зрения статья демонстрирует применимость предлагаемого подхода к российским данным, что открывает перспективы для снижения информационной асимметрии на рынке юридических услуг и повышения качества защиты прав подсудимых.

Статья имеет следующую структуру. Ниже представлен обзор литературы, охватывающий существующие подходы к оценке качества адвокатов, методы прогнозирования судебных решений, аналогии из спортивной аналитики и медицины, эмпирические исследования российского уголовного правосудия, а также рейтинговые системы и байесовские подходы. Следующий раздел содержит детальное описание предлагаемой методологии, включая данные и источники информации, трёхэтапную модель прогнозирования, применяемые методы машинного обучения, формулы расчёта рейтинга и учёт специфических факторов. Четвёртый раздел обсуждает ограничения методологии и перспективные направления исследований. В заключении приведены наиболее важные результаты работы.

Обзор литературы

Оценка качества работы адвокатов: существующие подходы

Систематическое изучение вопроса о том, влияет ли (и в какой мере) качество работы адвоката на исходы судебных дел, имеет длительную историю в правовой и экономической науке. Самая известная работа в данной области принадлежит М. Галантеру, который разграничил участников процесса, обращающихся в суд единожды (так называемые

«one-shotters») и регулярно участвующих в судебных разбирательствах, накапливающих экспертизу специалистов («повторяющихся игроков» — игроков с опытом, или «repeat players») [Galanter, 1974]. Галантер показал, что опытные игроки (это, как правило, компании, обладающие более квалифицированными юристами, но не всегда) систематически добиваются лучших результатов благодаря аккумулированным знаниям, стратегическому планированию и институциональным преимуществам. Эта теоретическая рамка до сих пор остаётся наиболее цитируемой основой для понимания структурного неравенства в доступе к качественной юридической помощи. Существенным выводом работы было то, что она показала, — важен именно опыт участия в судебных делах, а не только и не столько «ресурсы», за счёт которых большие фирмы могут побеждать маленьких игроков.

Одной из немногих эмпирических работ, использующих случайное распределение дел для идентификации эффекта адвоката, является исследование Абрамса и Юна, выполненное на данных публичных защитников Лас-Вегаса [Abrams, Yoon, 2007]. Авторы обнаружили значительную гетерогенность в результативности адвокатов: разница между 90-м и 10-м процентилем составила 14 п. п. в вероятности лишения свободы подзащитного. При этом опыт работы оказался значимым предиктором качества, тогда как престижность юридической школы не показала значимости. Аналогичная стратегия идентификации на основе случайного распределения дел была применена Васконселосом и соавторами в контексте бразильской гражданской юстиции: на выборке из 30 821 решения по 386 адвокатам авторы обнаружили статистически значимые различия в результативности примерно в 43% исследованных команд [Vasconcelos, Oliveira, Netto, 2018]. Примечательно, что в гражданских делах корреляция между опытом и результативностью не была обнаружена, что указывает на контекстуальную специфику качества юридического образования. К сожалению, подобных работ с каузальной идентификацией «эффекта адвоката» крайне мало, поскольку в реальной практике судебные дела почти никогда не распределяются между адвокатами полностью случайно (в этом отношении интересно случайное распределение дел «по назначению», введённое в России, которое, однако, сложно использовать из-за отсутствия в текстах судебных решений информации о типе контракта адвоката).

На уровне апелляционного судопроизводства Нельсон и Эпштейн продемонстрировали, что опытные адвокаты в Верховном суде США значительно чаще выигрывают дела и получают более благоприятные приговоры по сравнению с адвокатами, впервые представляющими дела в высшей инстанции [Nelson, Epstein, 2022]. Лазарус зафиксировал формирование элитного частного «клуба юристов» Верховного суда, представители которого непропорционально доминируют по результативности [Lazarus, 2008]. Эти исследования подтверждают, что эффект адвоката системен и обусловлен имеющимися институтами, но его нельзя свести просто к качеству юридического образования или к опыту.

Честерман, проведя собственное эмпирическое исследование в Верховном суде Сингапура, обнаружил, что более крупные и обеспеченные ресурсами юридические фирмы достигают лучших по сравнению со средними результатов. Однако более опытные индивидуальные адвокаты иногда демонстрируют более низкие показатели успеха. Предположительно это — следствие самоотбора в более сложные дела [Chesterman, 2020]. Такой вывод позволяет сформулировать важнейшую методологическую проблему: без контроля сложности дел «чистые» показатели результативности (например, доля побед) могут быть контринтуитивными и вводящими в заблуждение.

Другим важным методологическим ориентиром для нас стала работа Грейнера и Паттанаяк, проведших первое рандомизированное контролируемое испытание эффекта юридического представительства в апелляциях по страхованию от безработицы [Greiner, Pattanayak, 2012]. Авторы не обнаружили статистически значимого эффекта на исходы дел. Неожиданный нулевой результат, однако, подчеркивает необходимость строгой методологической базы для исследований эффективности адвокатов и невозможность

экстраполяции результатов между разными контекстами. Систематический обзор предшествующей литературы, выполненный в этой работе, показал, что подавляющее большинство исследований о влиянии юристов на исходы дел страдает от столь серьёзных методологических проблем, что их выводы не могут считаться надёжными.

В последние годы появились подходы, использующие большие данные и алгоритмические методы для оценки качества юридических услуг. Можон, Махари и Лера предложили модель для попарного сравнения юридических фирм, рассматривая каждый гражданский процесс как соревнование между фирмами [Mojon, Mahari, Lera, 2025]. На выборке из более чем 60 540 гражданских дел в США авторы показали, что существующие репутационные рейтинги (Vault, ALM) слабо коррелируют с реальной результативностью, тогда как основанные на реальных данных рейтинги повышают прогностическую точность примерно на 10% по сравнению с базовым уровнем. Этот результат наглядно демонстрирует разрыв между восприятием качества (разными субъективными рейтингами) и объективными результатами адвокатов на рынке юридических услуг.

Прогнозирование судебных решений методами машинного обучения

Второе направление литературы, критически важное для предлагаемой методологии, связано с прогнозированием судебных решений с помощью методов машинного обучения. Способность алгоритмов предсказывать исходы дел на основе их характеристик — необходимое условие для построения value-added моделей качества работы адвокатов: чем точнее прогноз, тем более информативно остаточное отклонение, атрибутируемое действиям защитника.

Ключевой работой в данном направлении стало исследование Алетраса с соавторами, применивших методы обработки естественного языка (NLP) к прогнозированию решений Европейского суда по правам человека [Aletras et al., 2016]. Авторы достигли средней точности 79% при бинарной классификации (нарушение или его отсутствие). Примечательно, что раздел фактов по делу оказался наиболее сильным предиктором — результат, согласующийся с теорией правового реализма.

Катц, Боммарито и Блэкман разработали эволюционирующий во времени классификатор Random Forest для прогнозирования решений Верховного суда США. Базой исследования были более чем 240 000 голосов судей и 28 000 исходов дел за период 1816–2015 гг. Они достигли 70,2% точности на уровне анализа дел при использовании исключительно тех, данные которых были доступны до судебного решения [Katz, Bommarito, Blackman, 2017]. Это исследование продемонстрировало, что данные методы существенно превосходят линейные модели в задачах судебного прогнозирования.

В области прогнозирования рецидивизма — задачи, структурно близкой к прогнозированию судебных решений, — накоплен значительный объём сравнительных исследований. Толленаар и Ван дер Хейден систематически сопоставили градиентный бустинг (XGBoost-типа), Random Forest и классическую логистическую регрессию для прогнозирования рецидивизма на нидерландских и северокаролинских данных [Tollenaar, van der Heijden, 2019]. Для бинарного прогноза методы машинного обучения не показали значимого улучшения по сравнению с хорошо специфицированной логистической регрессией, хотя градиентный бустинг продемонстрировал преимущество в моделях с временным компонентом. Дрессель и Фарид показали, что широко используемый алгоритм COMPAS со 137 признаками не превосходит по точности (примерно 65%) простую двухфакторную логистическую регрессию [Dressel, Farid, 2018], ставя под вопрос целесообразность усложнения алгоритмов в отдельных контекстах.

Систематические обзоры методов машинного обучения в прогнозировании судебных решений подтверждают общую картину: большинство методов достигают точности выше 70% при бинарной классификации, а наиболее распространёнными алгоритмами являются

SVM, Random Forest, нейронные сети и наивный байесовский классификатор [Rosili et al., 2021]. Обзор Травайни и его соавторов, выполненный по стандарту PRISMA на двенадцати высококачественных исследованиях прогнозирования рецидивизма, сообщает о средней точности 0,81 и AUC 0,74 [Travaini et al., 2022]. Важно, что сложность модели не обеспечивает её устойчивого превосходства над более простыми базовыми линейными регрессиями, что подчёркивает значимость тщательного подбора метода под конкретную задачу.

Отдельного внимания заслуживает работа Менезес-Нето и Клементино, обучивших три модели глубокого обучения (ULMFiT, BERT, Big Bird) на более чем 612 000 апелляций бразильских федеральных судов [Menezes-Neto, Clementino, 2022]. Все модели превзошли 22 высококвалифицированных юридических экспертов в прогнозировании исходов апелляций (лучшая модель: MCC = 0,37 против 0,13 у экспертов). Этот результат непосредственно подтверждает, что алгоритмические прогнозы судебных решений могут служить надёжной основой для построения value-added моделей.

Комплексный обзор NLP-приложений в юридической сфере, выполненный Ариаи с коллегами на основе 131 исследования, систематизирует методы анализа юридических текстов, включая прогнозирование решений, суммаризацию документов и извлечение именованных сущностей [Ariai, Mackenzie, Demartini, 2025]. Авторы выделяют 16 открытых исследовательских проблем, в том числе обнаружение смещений, интерпретируемость моделей и ограниченность открытых наборов данных. Русскоязычный систематический обзор международных исследований по прогнозированию судебных решений представлен в [Казун, 2024].

Аналогии из спортивной аналитики и медицины

Методологическое ядро предлагаемого подхода к оценке эффективности адвокатов имеет наиболее тесные параллели с двумя областями, в которых проблема измерения индивидуального вклада профессионала в условиях множества неконтролируемых факторов решается наиболее последовательно. Это спортивная аналитика и оценка работы специалистов в сфере медицины.

В спортивной аналитике центральная задача — оценка индивидуального вклада игрока в командный результат. Метрика Adjusted Plus-Minus (APM) и её регуляризованные варианты (Regularized APM, RAPM) решают эту задачу с помощью регрессионных моделей, контролирующих силу партнёров и соперников. Макдональд адаптировал методологию APM/RAPM, применявшуюся в оценках баскетболистов, для игроков в хоккей, используя гребневую регрессию для получения стабильных оценок наступательного и оборонительного вклада каждого игрока [Macdonald, 2012]. Регуляризация выступает ключевым техническим решением, позволяющим изолировать индивидуальный вклад при ограниченных данных. Это проблема, аналогичная задаче по оценке работы адвоката с малым числом дел. Коллектив авторов под руководством Грассетти расширил модель RAPM для баскетбола, продемонстрировав продвинутое техники регуляризации при оценке индивидуального вклада в командной игре [Grassetti et al., 2021].

Вместе с тем Гимире, Эрлих и Сандерс показали на данных NBA за 2013–2019 гг., что стандартные APM-модели не полностью учитывают командную игру: оценки RPM зависят от качества партнёров на площадке [Ghimire, Ehrlich, Sanders, 2020]. А Брэйв, Баттерс и Робертс показали, что примерно 40% необъяснённой вариации в результатах команд MLB связано с синергетическими эффектами командной работы игроков [Brave, Butters, Roberts, 2019]. Эти результаты наталкивают на возможности решения аналогичных задач в юридическом контексте: изолирование индивидуального вклада в командной игре имеет принципиальные методологические ограничения.

Метрика Wins Above Replacement (WAR) в бейсболе предлагает иную, но комплементарную логику: она измеряет, насколько побед данный игрок превосходит «replacement-level» — легкодоступного «заменителя» [Baumer, Jensen, Matthews, 2015]. Концепция

replacement level применима к оценке адвокатов, позволяя оценить, насколько данный защитник превосходит типичного («среднего») адвоката по данному типу дел. В футболе метрика Expected Goals (xG) даёт оценку, исходящую не от прямых результатов, а ориентированную на зависимость от скорректированной вероятности успеха [Anzer, Bauer, 2021].

В медицинской литературе оценка результативности работы индивидуальных хирургов представляется как наиболее прямая аналогия с задачей оценки адвокатов. Холл и его соавторы на данных реестра ACS NSQIP (39 976 случаев, 197 хирургов, 9 больниц) построили иерархические регрессионные модели с корректировкой на риск для оценки индивидуальных показателей хирургов по возможности осложнений (морбидности), раневой инфекции и смертности [Hall et al., 2015]. Авторы ввели метрики надёжности принципа «сигнал-к-шуму» (signal-to-noise reliability) для определения условий, при которых индивидуальные профили хирургов могут статистически различаться, и показали, что частота возникновения у пациента осложнений лучше различима, чем смертность. Эта методология непосредственно переносима на задачу профилирования адвокатов: она определяет минимальные требования к статистической мощности, методы корректировки на риск и учёт институционального контекста.

Крелл и соавторы, используя данные ACS NSQIP, установили, что большинство стандартных показателей качества имеют низкую надёжность для различения больниц — в особенности для малоинвазивных специальностей или редких событий [Krell et al., 2014]. Дюкло и коллеги разработали графические контрольные карты для мониторинга индивидуальных хирургов, скорректированные как на характеристики пациентов, так и на опыт хирурга [Maruthappu et al., 2014]. Питчес, Мохаммед и Лилфорд в систематическом обзоре 36 исследований обнаружили, что корреляция между скорректированной на риски смертностью и реальным качеством медицинской помощи неустойчива: лишь около половины исследований подтвердили тут положительную связь [Pitches, Mohammed, Lilford, 2007]. Этот результат может отражать неизмеренные факторы сложности случаев, а не подлинные различия в качестве. Чен с коллегами предложили композитные показатели качества, комбинирующие различные факторы [Chen et al., 2013], — методологический подход, релевантный для построения многомерных рейтинговых систем адвокатов.

Эмпирические исследования российского уголовного правосудия

Контекст российского уголовного правосудия обладает рядом специфических характеристик, которые существенно влияют на проектирование методологии оценки эффективности адвокатов. Центральной особенностью отечественной ситуации является выраженный обвинительный уклон: по данным исследований, доля оправдательных приговоров составляет менее 0,2%, что означает осуждение более чем 499 из каждых 500 подсудимых, прошедших через судебный процесс [Панеях, 2012]. Э. Панеях продемонстрировала, что данный уклон обусловлен институциональным давлением на судей, включающим неформальную иерархическую подотчётность и показатели эффективности, привязанные к низким показателям оправданий.

В. Волков провёл первый масштабный количественный анализ более 5 млн приговоров российских уголовных судов, зафиксировав систематические диспропорции в назначении наказаний по полу, занятости, гражданству и должностному положению [Волков, 2014]. Эти результаты свидетельствуют о наличии как легальных, так и экстралегальных факторов, формирующих судебные решения, что делает необходимым их учёт в прогностических моделях. В последующей работе Волков, используя более 350 тыс. судебных решений, изолировал каузальный эффект присутствия прокурора: в делах с участием государственного обвинителя вероятность оправдания оказалась примерно в десять раз ниже [Volkov, 2021]. Дополнительные свидетельства систематических различий в назначении наказаний на материале дел об убийствах см. в [Zhuchkova, Kazun, 2023].

К. Титаев на случайной выборке из 10 тыс. решений районных судов за 2011 г. продемонстрировал, что содержание под стражей до суда значительно повышает вероятность последующего приговора к реальному лишению свободы, указывая на кумулятивный эффект ранних прокурорских решений на итоговые исходы [Titaev, 2017]. А. Виноградов, анализируя структурные причины обвинительного уклона на досудебной стадии, аргументировал, что основная системная причина тут — помещение следователей на сторону обвинения в структуре Уголовно-процессуального кодекса РФ [Виноградов, 2025].

Развитие вычислительной инфраструктуры для анализа российских правовых текстов представлено работой Д. Савельева и Р. Кучакова, создавших масштабный открытый корпус из 281 413 федеральных правовых текстов за 1991-2023 гг. (176 млн токенов с полной морфосинтаксической разметкой) [Saveliev, Kuchakov, 2024]. В рамках проекта «Разработка и апробация методики автоматизированного анализа текстов приговоров российских судов для социально-правовых исследований (на примере насильственных преступлений)» исследовательский коллектив НИУ ВШЭ собрал корпус из более 7 млн текстов судебных решений и разработал методику их анализа [Жучкова и др., 2025]. Эти корпуса текстов могут быть отправной точкой для решения задач, связанных с оценкой качества работы российских адвокатов.

Рейтинговые системы и байесовские подходы

Методология построения рейтингов профессионалов на основе результатов их деятельности тесно связана с классической теорией парных сравнений. Модель Брэдли-Терри, предложенная в 1952 г., назначает каждому объекту латентный параметр силы, определяющий вероятность победы в парном сравнении [Bradley, Terry, 1952]. М. Каттелан в обширном обзоре моделей парных сравнений систематизировала расширения базовой модели для учёта ковариат, ничьих, порядковых эффектов и зависимых наблюдений [Cattelan, 2012]. Эти методы непосредственно применимы к конструированию рейтингов адвокатов на основе исходов дел.

Отдельно следует отметить опыт таких систем, как TrueSkill. Мастерство каждого игрока представляется в них как гауссовское распределение, учитывающее и среднее значение его навыка, и неопределённость, связанную с его оценкой [Herbrich, Minka, Graepel, 2006]. Этот подход можно использовать для оценки качества работы адвокатов с небольшим числом дел.

Фундаментальным инструментом для работы с малыми выборками в задачах ранжирования является эмпирическая байесовская сшивка (shrinkage). Эфрон и Моррис в классической работе по оценке «бейсбольных средних» показали, что объединение информации между объектами и сжатие индивидуальных оценок к общему среднему снижает общую ошибку предсказания [Efron, Morris, 1975]. Этот принцип лежит в основе сглаживающих процедур в методологии оценки работы адвокатов: при малом числе дел индивидуальная оценка эффективности «подтягивается» к средней по совокупности, предотвращая экстремальные рейтинги на основе недостаточных данных.

Методы матричной факторизации, получившие известность в контексте рекомендательных систем и Netflix Prize, предоставляют дополнительный инструмент для моделирования латентных факторов качества [Koren, Bell, Volinsky, 2009]. Хотя нельзя не отметить, что непосредственное применение коллаборативной фильтрации к задаче оценки адвокатов не является прямым, лежащая в её основе идея декомпозиции наблюдаемых результатов на латентные факторы концептуально близка к предлагаемому подходу.

Вопрос о так называемой «алгоритмической справедливости» очень важен в контексте обсуждений о работе алгоритма COMPAS, предсказывающего в США вероятность преступного рецидива. Дело в том, что алгоритм завышал риски для чернокожих подсудимых [Angwin et al., 2016], определив на основе обучающих данных, что этот фактор — значимый

предиктор. Карими-Хагиги и Кастильо предложили метод смягчения подобных алгоритмических смещений [Karimi-Haghighi, Castillo, 2021]. Эти работы подчёркивают важность контроля за подобными систематическими смещениями, которые модели могут «наследовать» из-за смещений в данных реального мира.

Методология

Настоящий раздел содержит детальное описание предлагаемой методологии оценки эффективности работы адвокатов по уголовным делам. Она основана на принципе value-added: прогностические модели машинного обучения строят ожидаемые исходы дел на основе характеристик, не зависящих от действий защитника, а систематическое отклонение фактических исходов от предсказанных рассматривается как вклад адвоката.

Данные и источники информации

Информационной основой методологии служат тексты судебных решений по уголовным делам, публикуемые в открытом доступе на сайтах судов общей юрисдикции в соответствии с Федеральным законом № 262-ФЗ от 22 декабря 2008 г. «Об обеспечении доступа к информации о деятельности судов в Российской Федерации». Суммарный объём анализируемых решений составляет более 7 млн документов за период 2010 — 2025 гг. Из текстов решений методами автоматизированного извлечения информации можно получить следующие переменные:

- ▶ характеристики преступления: статья Уголовного кодекса Российской Федерации (с точностью до части статьи), категория тяжести преступления;
- ▶ характеристики подсудимого: наличие рецидива, смягчающие обстоятельства (явка с повинной, активное содействие расследованию, наличие малолетних детей и др.), отягчающие обстоятельства (совершение преступления в группе, состояние опьянения и др.);
- ▶ институциональные характеристики: идентификатор суда, регион, социально-демографические характеристики населённого пункта (численность населения, уровень образования, структура занятости);
- ▶ временные характеристики: год и месяц вынесения решения;
- ▶ процессуальные характеристики: рассмотрение дела в особом порядке, мера пресечения;
- ▶ идентификация адвоката: ФИО защитника, установленные путём сопоставления с данными Единого государственного реестра адвокатов Министерства юстиции Российской Федерации [Минюст России, 2026].

Принципиальным элементом конструирования предикторного пространства является разделение переменных на две группы: а) факторы, не зависящие от действий адвоката (состав преступления, рецидив, регион, суд, время); б) факторы, на которые адвокат потенциально может повлиять (процессуальная форма рассмотрения дела, избранная мера пресечения, а также непосредственно исходы дел). Прогностические модели обучаются исключительно на переменных первой группы, что обеспечивает экзогенность прогнозов относительно действий защитника и позволяет интерпретировать остаточное отклонение как вклад адвоката.

Этапы прогнозирования судебных решений

Прогнозирование исходов судебных решений осуществляется поэтапно, что обусловлено логикой движения уголовного дела в суде и наличием нескольких ключевых процессуальных развилок, каждая из которых формирует самостоятельный объект анализа. Трёхэтапная структура обеспечивает методологическую корректность сопоставления прогнозируемых и фактических результатов на каждом уровне принятия решения.

Этап 1. Прогнозирование наличия судимости (судимость/несудимость). На первом этапе оценивается вероятность того, что итоговый судебный исход повлечёт судимость. К категории «несудимость» относятся прекращение дела, возвращение дела прокурору, примирение сторон, судебный штраф, оправдательный приговор и иные исходы, не влекущие судимости. К категории «судимость» относятся обвинительные приговоры, завершающиеся её возникновением (в том числе дела, которые закончились условным сроком). С учётом того, что, как было отмечено выше, оправдательные приговоры в российских судах чрезвычайно редки, основная эмпирическая граница на данном этапе проходит именно между исходами, влекущими и не влекущими судимость. Модели на данном этапе обучаются на полной совокупности уголовных дел.

Этап 2. Прогнозирование вида наказания при вынесении приговора (реальное лишение свободы или альтернативные наказания). Внутри категории дел, завершившихся приговором, прогнозируется вероятность назначения реального срока лишения свободы по сравнению со всеми альтернативными исходами (условное осуждение, штраф, обязательные работы и др.). Этот этап представляет собой ключевой момент оценки тяжести последствий для подсудимого. Модели обучаются на подвыборке дел, по которым был вынесен приговор, что исключает влияние структурных особенностей дел, завершившихся постановлением о прекращении дела.

Этап 3. Прогнозирование продолжительности срока реального лишения свободы. Для дел, по которым было назначено реальное лишение свободы, прогнозируется продолжительность срока (в месяцах). Данный этап направлен на выявление способности адвоката снизить тяжесть наказания в тех случаях, когда назначение реального срока неизбежно. Модели обучаются исключительно на подвыборке дел с назначенным реальным сроком.

Совокупное применение трёх этапов прогнозирования обеспечивает построение детализированной структуры ожидаемых исходов, позволяющей оценить вклад адвоката на каждом уровне процессуального решения: от вопроса «влечёт ли итоговый исход судимость?» через вопрос «назначено ли реальное лишение свободы?» до вопроса «каков срок?».

Методы машинного обучения

Основными моделями, применяемыми в рамках предлагаемой нами методологии, являются ансамблевые методы Random Forest и Extreme Gradient Boosting (XGBoost), а также обобщённые линейные модели (GLM) в качестве базовой модели для сравнения, что соответствует стандартным международным практикам.

Random Forest применяется для построения совокупности деревьев решений на основе случайных подвыборок предикторов и наблюдений (bagging + random feature selection), что обеспечивает минимизацию риска переобучения и высокую стабильность прогнозов [Katz, Bommarito, Blackman, 2017]. Метод доказал свою эффективность в задачах судебного прогнозирования: Катц с соавторами продемонстрировали устойчивое превосходство Random Forest при прогнозировании решений Верховного суда США.

Extreme Gradient Boosting (XGBoost) основан на процедуре последовательного улучшения качества прогнозов путём минимизации функции потерь на каждом шаге и демонстрирует высокую точность при наличии сложных взаимодействий между переменными [Tollenaar, van der Heijden, 2019]. В задачах прогнозирования бинарных исходов в уголовном правосудии XGBoost неоднократно демонстрировал конкурентные результаты по сравнению с другими методами.

Обобщённые линейные модели (GLM) рассматриваются в качестве базового метода для сопоставления результатов. Предварительные испытания показывают, что показатели точности GLM систематически уступают ансамблевым методам на большинстве подвыборок данных, что согласуется с результатами сравнительных исследований [Dressel, Farid, 2018; Tollenaar, van der Heijden, 2019].

Для оценки качества прогнозов применяется дифференцированная система метрик. На этапах 1 и 2 (бинарная классификация) используется показатель AUC (Area Under the ROC Curve), позволяющий оценить способность модели различать благоприятные и неблагоприятные исходы. На этапе 3 (непрерывная переменная — срок в месяцах) используется RMSE (Root Mean Square Error), измеряющий среднеквадратичное отклонение прогнозов от фактических значений.

Оценка качества проводится посредством пятиблочной кросс-валидации, обеспечивающей объективность и статистическую надёжность. По итогам обучения для каждой подвыборки и этапа выбирается модель с наилучшими показателями качества.

Методика расчёта рейтинга эффективности адвокатов

Расчёт рейтинга эффективности адвокатов осуществляется в ходе нескольких последовательных шагов, начиная от «сырой» разницы между прогнозом и фактом и заканчивая итоговым децильным рейтингом.

Шаг 1. Расчёт «сырой» дельты. Для каждого дела k на каждой стадии s рассчитывается разница между фактическим и предсказанным исходом по формулам (1a) и (1b):

Для бинарных исходов (этапы 1 и 2):

$$\delta_{k,s} = \gamma_{k,s} - \rho_{k,s}, \quad (1a)$$

где $\gamma_{k,s} \in \{0,1\}$ — фактический результат; $\rho_{k,s}$ — предсказанная наилучшей моделью вероятность исхода.

Для непрерывной переменной (этап 3 — продолжительность лишения свободы в месяцах):

$$\delta_{k,s} = \hat{x}_{k,s} - x_{k,s}, \quad (1b)$$

где $\hat{x}_{k,s}$ — наилучшее предсказание модели; $x_{k,s}$ — фактическое значение переменной.

В обеих формулах положительная дельта означает более благоприятный для подсудимого исход, чем предсказывала модель. В формуле (1a) положительное значение возникает, когда фактический исход лучше ожидаемого (например, дело прекращено, хотя прогнозировался приговор). В формуле (1b) положительное значение означает, что фактический срок оказался короче предсказанного.

Шаг 2. Штраф за последовательность этапов. Если на предыдущей стадии ($s-1$) была предсказана высокая вероятность благоприятного исхода $\rho_{k,s-1}$, но фактически он не был достигнут, то на стадии s адвокату назначается штраф $\phi_{k,s-1}$ по формуле (2):

$$\phi_{k,s-1} = \{0, \text{ если } \gamma_{k,s-1} = 1 \text{ или } \rho_{k,s-1} \leq \bar{\rho}; \lambda \cdot \frac{1}{1-\bar{\rho}} (\rho_{k,s-1} - \bar{\rho}), \text{ иначе}\}, \quad (2)$$

где λ — максимально возможный штраф (параметр; $\lambda = 0,2$); $\bar{\rho}$ — порог вероятности благоприятного исхода ($\bar{\rho} = 0,6$). Данный штраф реализует принцип последовательности: если модель предсказывала, что дело с высокой вероятностью завершится прекращением, но этого не произошло, то бонус адвоката за относительно мягкий приговор на следующей стадии уменьшается.

Штрафы применяются кумулятивно:

$$\delta_{k,s} = \delta_{k,s} - |\delta_{k,s}| \cdot \sum_{t < s} \phi_{k,t}, \quad (3)$$

где левая часть формулы обозначает дельту после применения штрафов за последовательность этапов.

Шаг 3. Корректировка на командную работу. Если в деле участвуют несколько адвокатов, то индивидуальный вклад каждого из них уменьшается по формуле (4):

$$\tilde{\delta}_{k,s} = \dot{\delta}_{k,s} - |\dot{\delta}_{k,s}| \cdot \min\{\mu, n_k - 1\} \cdot \vartheta, \quad (4)$$

где ϑ — размер штрафа за одного дополнительного адвоката ($\vartheta = 0,05$); μ — максимальное число дополнительных адвокатов, за которых накладывается штраф ($\mu = 5$); n_k — общее количество адвокатов в деле k . Данная корректировка отражает принцип, аналогичный учёту командных эффектов в спортивной аналитике [Ghimire, Ehrlich, Sanders, 2020; Brave, Butters, Roberts, 2019]: при коллективном достижении результата индивидуальная атрибуция снижается.

Шаг 4. Агрегирование по делам. Пусть $\Omega_{i,s}$ обозначает множество дел, в которых адвокат i участвовал на стадии s . Средняя дельта адвоката i на стадии s рассчитывается по формуле:

$$\delta_{i,s}^{avg} = \frac{1}{N_{i,s}} \sum_{k \in \Omega_{i,s}} \tilde{\delta}_{k,s}, \quad (5)$$

где $N_{i,s}$ — количество дел адвоката i на стадии s .

Шаг 5. Байесовское сглаживание (shrinkage). При малом количестве дел индивидуальная средняя дельта может оказаться нестабильной. Для адвокатов с $N_{i,s} < \alpha_s$ (где α_s — минимальное число дел для устойчивой оценки; для всех стадий $\alpha_s = 10$) применяется сглаживание:

$$\hat{\Delta}_{i,s} = \begin{cases} \delta_{i,s}^{avg}, & \text{если } N_{i,s} \geq \alpha_s \\ \frac{N_{i,s}}{\alpha_s} \cdot \delta_{i,s}^{avg} + \frac{\alpha_s - N_{i,s}}{\alpha_s} \cdot \bar{\Delta}_s, & \text{иначе} \end{cases}, \quad (6)$$

где левая часть формулы (6) представляет собой итоговую сглаженную среднюю дельту адвоката i на стадии s , тогда как $\bar{\Delta}_s$, т. е. среднее значение дельты на стадии s среди всех адвокатов, выступает в качестве shrinkage-ориентира. Данная процедура реализует принцип эмпирической байесовской shrinkage-оценки, обоснованной в классических работах по статистике малых выборок [Efron, Morris, 1975]: при недостаточном количестве наблюдений индивидуальная оценка «подтягивается» к средней, предотвращая экстремальные рейтинги. По мере накопления данных ($N_{i,s} \rightarrow \alpha_s$) вес глобального среднего убывает до нуля, и оценка полностью определяется индивидуальными результатами.

Шаг 6. Позиция в распределении (ECDF). Для множества всех адвокатов на стадии s , обозначаемого D_s , пусть $m_s = |D_s|$. Для каждого адвоката i вычисляется значение эмпирической функции распределения по формуле (7):

$$T_{i,s} = \frac{1}{m_s} \sum_j I(\hat{\Delta}_{j,s} \leq \hat{\Delta}_{i,s}), \quad (7)$$

где $I(\cdot)$ — индикаторная функция, принимающая значение 1, когда условие в скобках выполняется, и 0 в противном случае.

Шаг 7. Агрегирование по стадиям. Для получения итоговой оценки адвоката по всем стадиям значения $T_{i,s}$ взвешиваются на количество дел на каждой стадии. Пусть p_i — множество стадий, в которых адвокат участвовал. Тогда итоговая оценка рассчитывается по формуле (8):

$$\bar{T}_i = \frac{1}{\sum_{s \in p_i} N_{i,s}} \sum_{s \in p_i} N_{i,s} \cdot T_{i,s}. \quad (8)$$

Шаг 8. Итоговая ECDF. По множеству итоговых агрегированных значений рассчитывается финальная позиция каждого адвоката в общем распределении по формуле (9):

$$r_i = \frac{1}{|D|} \sum_j I(\bar{T}_j \leq \bar{T}_i). \quad (9)$$

Шаг 9. Децильный рейтинг. Итоговый рейтинг адвоката определяется его позицией в соответствующем дециле по формуле (10):

$$R_i = q, \text{ если } \frac{q-1}{10} < r_i \leq \frac{q}{10}, \quad q = 1, \dots, 10 \quad (10)$$

Значение R_i представляет собой итоговый децильный рейтинг эффективности адвоката i : $R_i = 10$ соответствует высшему децилю (наиболее эффективные адвокаты), $R_i = 1$ — низшему. Рейтинг рассчитывается отдельно по каждой статье Уголовного кодекса, что обеспечивает специализированную оценку в рамках каждого типа дел.

Учёт специфических факторов

Предлагаемая методология включает ряд механизмов, направленных на учёт специфических факторов, влияющих на корректность оценки.

Командные эффекты. Формула (4) реализует дисконтирование индивидуального вклада при участии в деле нескольких адвокатов. Данный механизм концептуально аналогичен учёту командных эффектов в спортивной аналитике: так же, как метрики АРМ в командных видах спорта корректируют индивидуальные оценки с учётом вклада партнёров [Macdonald, 2012; Grassetti et al., 2021], формула (4) снижает атрибуцию индивидуальному адвокату при совместной защите. Параметры $\vartheta = 0,05$ и $\mu = 5$ обеспечивают умеренное дисконтирование: при двух адвокатах каждый получает 95% от полной дельты; при шести и более — 75%.

Сделки с правосудием (особый порядок). Значительная часть уголовных дел в России рассматривается в особом порядке, что существенно ограничивает пространство возможных исходов и, соответственно, возможность атрибуции результата качеству защиты. Поскольку выбор особого порядка может быть связан со стратегией защиты, в базовой спецификации этот показатель не включается в набор экзогенных предикторов; вместо этого он используется для стратификации выборки и проверки устойчивости оценок.

Малые выборки. Формула (6) реализует байесовское сглаживание, направленное на повышение устойчивости рейтинговых оценок при малом числе дел. Данный подход непосредственно соответствует методологии, описанной в классической работе по оценке бейсбольных средних [Efron, Morris, 1975], а также в практике медицинских регистров, где аналогичная проблема решается с помощью иерархических моделей с корректировкой на объём наблюдений [Hall et al., 2015; Krell et al., 2014]. Параметр $\alpha_s = 10$ означает, что для адвокатов с числом дел на данной стадии менее 10 рейтинг приближается к среднему значению, т.е. к центру шкалы (5–6 из 10).

Последовательность этапов. Формулы (2) и (3) реализуют принцип последовательной атрибуции: если на предшествующей стадии была высока вероятность благоприятного исхода, который тем не менее не был реализован, то бонус на последующей стадии уменьшается. Это предотвращает ситуацию, при которой адвокат, «потерявший» благоприятный исход на ранней стадии (например, модель предсказывала прекращение дела, но дело дошло до приговора), получает непропорционально высокий бонус за незначительное смягчение приговора.

Социально-демографические факторы. Включение социально-демографических характеристик населенного пункта (численность, уровень образования, структура

занятости) в предикторное пространство моделей обеспечивает контроль межрегиональной вариации, обусловленной различиями в социальном контексте [Volkov, 2014]. Это позволяет корректнее сопоставлять эффективность адвокатов, работающих в различных субъектах Российской Федерации.

Ограничения и направления развития

Предлагаемая методология имеет ряд существенных ограничений, связанных как с характером доступных данных, так и с методологическими допущениями. Важно честно указать на них, хотя, как мы описали выше, предложенный подход позволяет аккуратно учесть многие особенности данных.

Во-первых, на основе имеющихся данных возможна оценка только судебной стадии. Однако важной, часто решающей, частью деятельности защитника является досудебная стадия: участие в следственных действиях, обжалование мер пресечения, работа с доказательной базой, сбор показаний свидетелей, психологическая поддержка доверителя. Многие высококвалифицированные адвокаты специализируются именно на досудебной защите и порой вообще не ходят в суд. Однако дела, не поступившие в суд, полностью выпадают из анализа. Это создаёт систематическое смещение: адвокаты, наиболее эффективные на досудебной стадии, могут быть недооценены или вообще не войти в рейтинговую систему. Таким образом, предложенный подход должен учитывать тот факт, что он позволяет оценивать только и исключительно ту часть работы адвоката, которая связана с работой на судебной стадии процесса. Это важно учитывать ещё и потому, что порой исходные условия, в которых адвокат начинает работать на судебной стадии, задаются неграмотной помощью другого адвоката на предыдущей стадии процесса, которую мы не можем наблюдать.

Во-вторых, в качестве стандартного ограничения предстаёт неполнота данных. Не все судебные решения размещаются в открытом доступе (некоторые не должны публиковаться в соответствии с законом, другие не появляются из-за технических или организационных проблем). Часть документов публикуется с существенными задержками (до одного–двух лет). Кроме того, из некоторых текстов судебных решений изымаются персональные данные участников процесса, включая ФИО адвокатов. Это означает, что даже при наличии текста судебного решения не всегда возможно точно соотнести его с конкретным адвокатом.

В-третьих, на данном этапе алгоритм не позволяет адекватно оценивать дела с несколькими подсудимыми. Предложенная методология не включает в расчёт рейтинга те дела, в которых одновременно представлены несколько подсудимых и/или одному подсудимому предъявлено несколько обвинений. Это ограничение обусловлено техническими сложностями автоматизированного извлечения информации из текстов приговоров отдельно для каждого обвиняемого: различные подсудимые могут иметь различные наборы смягчающих и отягчающих обстоятельств, один из них может признать вину, а другой нет, на этом основании они могут получить разные сроки. Безусловно, аккуратное чтение текстов часто позволяет чётко определять те или иные исходы и обстоятельства дела, но для машинной обработки миллионов текстов это оказывается препятствием. Впрочем, данное препятствие имеет исключительно техническую природу и при должном развитии алгоритмов в обозримом будущем будет устранено.

В-четвёртых, точная идентификация адвокатов на основе текстов решений не всегда возможна. Небольшое число адвокатов остаётся неидентифицированным вследствие наличия у них полных тёзок (совпадение фамилии, имени и отчества). Это встречается удивительно часто и может говорить, в том числе, о высоком уровне династичности данной профессии. В таких случаях невозможно достоверно связать конкретные дела с конкретным адвокатом. Дополнительно в реестре адвокатов встречаются чисто технические ошибки, которые могут препятствовать корректной привязке дел к защитникам.

В-пятых, важно учитывать проблему ненаблюдаемых факторов. Центральным методологическим допущением нашего подхода является то, что остаточное отклонение фактических исходов от предсказанных в значительной мере обусловлено качеством работы адвоката. Однако на это отклонение могут влиять и другие ненаблюдаемые факторы: сила доказательственной базы обвинения (не полностью отражённая в тексте приговора), личностные характеристики судьи или суда (которые не ловятся фиксированными эффектами), специфические обстоятельства дела, не зафиксированные в извлекаемых переменных. Эта проблема хорошо известна при решении аналогичных задач в медицинской литературе [Pitches, Mohammed, Lilford, 2007]. В целом она является общей для всех value-added моделей, включая оценку учителей [Chetty, Friedman, Rockoff, 2014a] и врачей [Hall et al., 2015].

В-шестых, может быть недооценён эффект самоотбора. Более опытные и квалифицированные адвокаты систематически берутся за более сложные дела, что может приводить к занижению их рейтинга при недостаточном контроле сложности [Chesterman, 2020]. Хотя включение в модели характеристик преступления и обстоятельств дела частично контролирует этот эффект, полностью избежать эффекта самоотбора без рандомизации распределения дел невозможно [Greiner, Pattanayak, 2012].

Кроме того, важно учитывать точность прогностических моделей. Качество рейтинговой системы непосредственно зависит от точности моделей, которые мы строим. Систематические обзоры показывают, что точность ML-моделей для бинарных судебных исходов обычно находится в диапазоне 70–81% [Travaini et al., 2022; Rosili et al., 2021], что означает наличие существенной доли необъяснённой дисперсии. Впрочем, важно отметить, что эта дисперсия и не должна быть нулевой, ведь внутри (помимо «шума» и прочих ненаблюдаемых факторов) лежит именно то, что нас интересует — качество работы адвоката.

Наконец, широкое применение предложенной методики должно учитывать этические соображения. Любая рейтинговая система, оценивающая профессионалов, сопряжена с рисками (в том числе из-за ограничений, описанных выше). Первый риск связан с возможным вредом для репутации честного адвоката. По ряду причин адвокат может быть несправедливо оценён вследствие ограничений данных (неполнота профиля, самоотбор в сложные дела и пр.). В связи с этим рейтинговая система не может и не должна использоваться в качестве единственного критерия выбора защитника и тем более как свидетельство ненадлежащего качества работы. Система не должна применяться для «очистки» профессиональных рядов и применения иных дисциплинарных мер. Второй этический риск связан с обсуждавшейся выше алгоритмической справедливостью. Систематические смещения в моделях (например, обусловленные расовыми, гендерными или социальными факторами) могут касаться и рейтинговых оценок [Angwin et al., 2016; Karimi-Haghighi, Castillo, 2021]. По этим причинам требуется аккуратная этическая экспертиза при любом прикладном применении данных алгоритмов.

Можно выделить некоторые перспективные направления развития предлагаемой методологии:

1. Расширение анализа на досудебную стадию. Интеграция данных о ходатайствах, обжаловании мер пресечения, результатов рассмотрения процессуальных вопросов на стадии следствия позволит более полно оценить вклад адвоката. Впрочем, в настоящее время такие данные не доступны публично.

2. Применение NLP-методов для анализа текстов. Использование различных языковых моделей (BERT, и их адаптаций для русского языка) для извлечения дополнительных признаков из текстов приговоров может повысить точность прогнозирования [Ariai, Mackenzie, Demartini, 2025; Menezes-Neto, Clementino, 2022; Saveliev, Kuchakov, 2024].

3. Динамические рейтинговые модели. Применение байесовских моделей, аналогичных TrueSkill [Herbrich, Minka, Graepel, 2006], позволит формировать динамические, обновляемые в реальном времени рейтинги с явным учётом неопределённости оценки.

4. Каузальная идентификация. Использование квазиэкспериментальных методов (инструментальные переменные, разрывный дизайн) на основе институциональных особенностей распределения дел может повысить надёжность каузальной интерпретации рейтинговых оценок [Abrams, Yoon, 2007].

5. Решение проблемы множественных подсудимых. Разработка алгоритмов автоматизированного извлечения информации отдельно для каждого подсудимого в групповых делах позволит существенно расширить выборку.

6. Композитные показатели. Полезно формирование многомерных рейтингов, комбинирующих исходы дел с процессными показателями (активность в представлении ходатайств, скорость работы), по аналогии с композитными показателями качества в медицине [Chen et al., 2013].

Заключение

Настоящая статья предлагает формализованную методологию оценки эффективности работы адвокатов по уголовным делам, основанную на сравнении прогнозируемых и фактических исходов судебных решений. Теоретической основой подхода является принцип value-added, успешно применяемый в экономике образования [Chetty, Friedman, Rockoff, 2014a; Chetty, Friedman, Rockoff, 2014b], медицине [Hall et al., 2015] и спортивной аналитике [Grassetti et al., 2021; Macdonald, 2012].

Наша методология включает несколько ключевых компонентов. Во-первых, трёх-этапную модель прогнозирования, отражающую логику последовательных процессуальных развилок в уголовном судопроизводстве: от вопроса о форме итогового решения (следует ли судимость) через вопрос о виде наказания (реальное лишение свободы или альтернативы) до прогнозирования продолжительности срока заключения. Во-вторых, алгоритм предполагает применение ансамблевых методов машинного обучения (Random Forest, XGBoost), продемонстрировавших эффективность в задачах судебного прогнозирования [Katz, Bommarito, Blackman, 2017]. В-третьих, методология включает систему корректировок, учитывающих последовательность этапов рассмотрения дела (формулы 2–3), командную работу нескольких адвокатов (формула 4) и нестабильность оценок при малом числе дел (формула 6). Последняя корректировка реализует принцип байесовского сглаживания, обоснованный для задач на малых выборках [Efron, Morris, 1975] и широко применяемый в медицинских исследованиях [Hall et al., 2015] и спортивной аналитике.

Предлагаемый подход, в случае его широкой реализации (при условии соблюдения этических ограничений), может быть использован для снижения информационной асимметрии на рынке юридических услуг в сфере уголовной защиты. Исследования показывают, что существующие репутационные рейтинги слабо коррелируют с реальной результативностью юридических фирм [Mojon, Mahari, Lera, 2025], а разброс в качестве работы адвокатов является значительным и имеет существенные последствия для клиентов [Abrams, Yoon, 2007; Nelson, Epstein, 2022]. Основанные на данных рейтинги предоставляют объективную, количественно обоснованную дополнительную информацию для участников рынка.

Вместе с тем методология имеет существенные ограничения, которые необходимо принимать во внимание. Ключевая проблема здесь — невозможность полного учёта ненаблюдаемых факторов [Pitches, Mohammed, Lilford, 2007]. Описанные выше ограничения данных также требуют осторожности в интерпретации результатов. Рейтинговая система предоставляет важную дополнительную информацию при выборе адвоката, но не может и не должна рассматриваться как единственный или безусловный критерий оценки. Впрочем, создание подобной системы, на наш взгляд, создало бы существенно больше позитивных эффектов для рынка юридических услуг (как в части «спроса», так и в части «предложения»), чем вызывало бы дополнительные сложности.

ЛИТЕРАТУРА

- Виноградов А.С. (2025). Обвинительный уклон в досудебном производстве по уголовным делам: причины и пути преодоления // *Журнал правовых и экономических исследований (Вестник ГИЭФ)*. Т. 59. № 1. С. 133–138. DOI: 10.26163/gief.2025.59.52.016.
- Волков В.В. (2014). *Социально-экономический статус и различия в вынесении приговоров: данные о рассмотрении уголовных дел в судах РФ* / Препринт Института проблем правоприменения при Европейском университете в Санкт-Петербурге. — СПб.: ИПП ЕУСПб.
- Жучкова С.В., Девятников В.Ю., Казун А.П., Белов М.Д., Сидорова О.И. (2025). Тексты судебных приговоров как источник данных для эмпирических исследований права в России // *Мониторинг общественного мнения: экономические и социальные перемены*. № 2. С. 170–192. DOI: 10.14515/monitoring.2025.2.2575.
- Казун А.П. (2024). Может ли искусственный интеллект прогнозировать решения суда? Систематический обзор международных исследований // *Мониторинг общественного мнения: экономические и социальные перемены*. № 5. С. 100–122. DOI: 10.14515/monitoring.2024.5.2580.
- Минюст России. (2026). Единый государственный реестр адвокатов. <https://minjust.gov.ru/ru/pages/advokaty-rossijskoj-federacii/> (дата обращения: 17.03.2026).
- Панях Э.Л. (2012). Практическая логика принятия судебных решений: дискреция под давлением и компромиссы за счёт подсудимого // *Как судьи принимают решения: эмпирические исследования права* / под ред. В.В. Волкова. — М.: Статут. С. 107–127.
- Abrams D.S., Yoon A.H. (2007). The Luck of the Draw: Using Random Case Assignment to Investigate Attorney Ability // *University of Chicago Law Review*. Vol. 74. No. 4. Pp. 1145–1177. DOI: 10.2307/20141859.
- Aletras N., Tsarapatsanis D., Preotiuc-Pietro D., Lampos V. (2016). Predicting judicial decisions of the European Court of Human Rights: a Natural Language Processing perspective // *PeerJ Computer Science*. Vol. 2. Art. e93. DOI: 10.7717/peerj-cs.93.
- Angwin J., Larson J., Mattu S., Kirchner L. (2016). Machine Bias: There's Software Used Across the Country to Predict Future Criminals. And It's Biased Against Blacks // *ProPublica*. May 23. <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing> (access date: 17.03.2026).
- Anzer G., Bauer P. (2021). A Goal Scoring Probability Model for Shots Based on Synchronized Positional and Event Data in Football (Soccer) // *Frontiers in Sports and Active Living*. Vol. 3. Art. 624475. DOI: 10.3389/fspor.2021.624475.
- Ariai F., Mackenzie J., Demartini G. (2025). Natural Language Processing for the Legal Domain: A Survey of Tasks, Datasets, Models, and Challenges // *ACM Computing Surveys*. Vol. 58. No. 6. Art. 163. Pp. 1–37. DOI: 10.1145/3777009.
- Baumer B.S., Jensen S.T., Matthews G.J. (2015). OpenWAR: An open source system for evaluating overall player performance in Major League Baseball // *Journal of Quantitative Analysis in Sports*. Vol. 11. No. 2. Pp. 69–84. DOI: 10.1515/jqas-2014-0098.
- Bradley R.A., Terry M.E. (1952). Rank Analysis of Incomplete Block Designs: The Method of Paired Comparisons // *Biometrika*. Vol. 39. No. 3–4. Pp. 324–345. DOI: 10.1093/biomet/39.3-4.324.
- Brave S.A., Butters R.A., Roberts K.A. (2019). Uncovering the sources of team synergy: Player complementarities in the production of wins // *Journal of Sports Analytics*. Vol. 5. No. 4. Pp. 247–279. DOI: 10.3233/JSA-190248.
- Cattelan M. (2012). Models for Paired Comparison Data: A Review with Emphasis on Dependent Data // *Statistical Science*. Vol. 27. No. 3. Pp. 412–433. DOI: 10.1214/12-STS396.
- Chen L.M., Staiger D.O., Birkmeyer J.D., Ryan A.M., Zhang W., Dimick J.B. (2013). Composite Quality Measures for Common Inpatient Medical Conditions // *Medical Care*. Vol. 51. No. 9. Pp. 832–837. DOI: 10.1097/MLR.0b013e31829fa92a.
- Chesterman S. (2020). Do Better Lawyers Win More Often? Measures of Advocate Quality and Their Impact in Singapore's Supreme Court // *Asian Journal of Comparative Law*. Vol. 15. No. 2. Pp. 250–280. DOI: 10.1017/asjcl.2020.13.
- Chetty R., Friedman J.N., Rockoff J.E. (2014a). Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates // *American Economic Review*. Vol. 104. No. 9. Pp. 2593–2632. DOI: 10.1257/aer.104.9.2593.
- Chetty R., Friedman J.N., Rockoff J.E. (2014b). Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood // *American Economic Review*. Vol. 104. No. 9. Pp. 2633–2679. DOI: 10.1257/aer.104.9.2633.
- Dressel J., Farid H. (2018). The accuracy, fairness, and limits of predicting recidivism // *Science Advances*. Vol. 4. No. 1. Art. eaao5580. DOI: 10.1126/sciadv.aao5580.
- Efron B., Morris C. (1975). Data Analysis Using Stein's Estimator and its Generalizations // *Journal of the American Statistical Association*. Vol. 70. No. 350. Pp. 311–319. DOI: 10.2307/2285814.
- Galanter M. (1974). Why the «Haves» Come Out Ahead: Speculations on the Limits of Legal Change // *Law & Society Review*. Vol. 9. No. 1. Pp. 95–160. DOI: 10.2307/3053023.
- Ghimire S., Ehrlich J.A., Sanders S.D. (2020). Measuring individual worker output in a complementary team setting: Does regularized adjusted plus minus isolate individual NBA player contributions? // *PLOS ONE*. Vol. 15. No. 8. Art. e0237920. DOI: 10.1371/journal.pone.0237920.

- Grassetti L., Bellio R., Di Gaspero L., Fonseca G., Vidoni P. (2021). An extended regularized adjusted plus-minus analysis for lineup management in basketball using play-by-play data // *IMA Journal of Management Mathematics*. Vol. 32. No. 4. Pp. 385–409. DOI: 10.1093/imaman/dpaa022.
- Greiner D.J., Pattanayak C.W. (2012). Randomized Evaluation in Legal Assistance: What Difference Does Representation (Offer and Actual Use) Make? // *Yale Law Journal*¹. Vol. 121. No. 8. Pp. 2118–2208.
- Hall B.L., Huffman K.M., Hamilton B.H., Paruch J.L., Zhou L., Richards K.E., Cohen M.E., Ko C.Y. (2015). Profiling Individual Surgeon Performance Using Information from a High-Quality Clinical Registry: Opportunities and Limitations // *Journal of the American College of Surgeons*. Vol. 221. No. 5. Pp. 901–913. DOI: 10.1016/j.jamcollsurg.2015.07.454.
- Herbrich R., Minka T., Graepel T. (2006). TrueSkill™: A Bayesian Skill Rating System // *Advances in Neural Information Processing Systems*. Vol. 19. Pp. 569–576. DOI: 10.5555/2976456.2976528.
- Karimi-Haghighi M., Castillo C. (2021). Enhancing a recidivism prediction tool with machine learning: effectiveness and algorithmic fairness // *Proceedings of the 18th International Conference on Artificial Intelligence and Law (ICAIL '21)*. Pp. 210–214. DOI: 10.1145/3462757.3466150.
- Katz D.M., Bommarito M.J. II, Blackman J. (2017). A general approach for predicting the behavior of the Supreme Court of the United States // *PLOS ONE*. Vol. 12. No. 4. Art. e0174698. DOI: 10.1371/journal.pone.0174698.
- Koren Y., Bell R., Volinsky C. (2009). Matrix Factorization Techniques for Recommender Systems // *Computer*. Vol. 42. No. 8. Pp. 30–37. DOI: 10.1109/MC.2009.263.
- Krell R.W., Hozain A., Kao L.S., Dimick J.B. (2014). Reliability of Risk-adjusted Outcomes for Profiling Hospital Surgical Quality // *JAMA Surgery*. Vol. 149. No. 5. Pp. 467–474. DOI: 10.1001/jamasurg.2013.4249.
- Lazarus R.J. (2008). Advocacy Matters Before and Within the Supreme Court: Transforming the Court by Transforming the Bar // *Georgetown Law Journal*. Vol. 96. No. 5. Pp. 1487–1564.
- Macdonald B. (2012). Adjusted Plus-Minus for NHL Players Using Ridge Regression with Goals, Shots, Fenwick, and Corsi // *Journal of Quantitative Analysis in Sports*. Vol. 8. No. 3. Pp. 1–24. DOI: 10.1515/1559-0410.1447.
- Maruthappu M., Carty M.J., Lipsitz S.R., Wright J., Orgill D., Duclos A. (2014). Patient- and surgeon-adjusted control charts for monitoring performance // *BMJ Open*. Vol. 4. No. 1. Art. e004046. DOI: 10.1136/bmjopen-2013-004046.
- Menezes-Neto E.J., Clementino M.B.M. (2022). Using deep learning to predict outcomes of legal appeals better than human experts: A study with data from Brazilian federal courts // *PLOS ONE*. Vol. 17. No. 7. Art. e0272287. DOI: 10.1371/journal.pone.0272287.
- Mojon A., Mahari R., Lera S.C. (2025). Data-driven law firm rankings to reduce information asymmetry in legal disputes // *Nature Computational Science*. Vol. 5. No. 11. Pp. 1010–1016. DOI: 10.1038/s43588-025-00899-2.
- Nelson M.J., Epstein L. (2022). Human Capital in Court: The Role of Attorney Experience in US Supreme Court Litigation // *Journal of Law and Courts*. Vol. 10. No. 1. Pp. 61–85. DOI: 10.1086/714577.
- Pitches D.W., Mohammed M.A., Lilford R.J. (2007). What is the empirical evidence that hospitals with higher risk-adjusted mortality rates provide poorer quality care? A systematic review of the literature // *BMC Health Services Research*. Vol. 7. Art. 91. DOI: 10.1186/1472-6963-7-91.
- Rosili N.A.K., Hassan R., Zakaria N.H., Kasim S., Che Rose F.Z., Sutikno T. (2021). A systematic literature review of machine learning methods in predicting court decisions // *IAES International Journal of Artificial Intelligence*. Vol. 10. No. 4. Pp. 1091–1102. DOI: 10.11591/ijai.v10.i4.pp1091-1102.
- Saveliev D., Kuchakov R. (2024). The Russian Legislative Corpus // *arXiv preprint arXiv:2406.04855*. 2024. DOI: 10.48550/arXiv.2406.04855.
- Titaev K. (2017). Pretrial detention in Russian criminal courts: a statistical analysis // *International Journal of Comparative and Applied Criminal Justice*. Vol. 41. No. 3. Pp. 145–161. DOI: 10.1080/01924036.2016.1239117.
- Tollenaar N., van der Heijden P.G.M. (2019). Optimizing predictive performance of criminal recidivism models using registration data with binary and survival outcomes // *PLOS ONE*. Vol. 14. No. 3. Art. e0213245. DOI: 10.1371/journal.pone.0213245.
- Travaini G.V., Pacchioni F., Bellumore S., Bosia M., De Micco F. (2022). Machine Learning and Criminal Justice: A Systematic Review of Advanced Methodology for Recidivism Risk Prediction // *International Journal of Environmental Research and Public Health*. Vol. 19. No. 17. Art. 10594. DOI: 10.3390/ijerph191710594.
- Vasconcelos C.C., Oliveira E.W., Netto W.L. (2018). The Impact of Attorneys on Judicial Decisions: Empirical Evidence from Civil Cases // *International Journal for Court Administration*. Vol. 9. No. 2. Pp. 32–42. DOI: 10.18352/IJCA.244.
- Volkov V. (2021). The prosecutor effect in trials for petty violent offences in Russia // *International Journal of Comparative and Applied Criminal Justice*. Vol. 45. No. 2. Pp. 207–220. DOI: 10.1080/01924036.2020.1732434.
- Zhuchkova S., Kazun A. (2023). Exploring Gender Bias in Homicide Sentencing: An Empirical Study of Russian Court Decisions Using Text Mining // *Homicide Studies*. Advance online publication. DOI: 10.1177/10887679231217159.

¹ Йельский университет (Yale University) (США) признан нежелательной организацией в России.

REFERENCES

- Abrams D.S., Yoon A.H. (2007). The Luck of the Draw: Using Random Case Assignment to Investigate Attorney Ability // *University of Chicago Law Review*. Vol. 74. No. 4. Pp. 1145–1177. DOI: 10.2307/20141859.
- Aletras N., Tsarapatsanis D., Preoŋiuc-Pietro D., Lamos V. (2016). Predicting judicial decisions of the European Court of Human Rights: a Natural Language Processing perspective // *PeerJ Computer Science*. Vol. 2. Art. e93. DOI: 10.7717/peerj-cs.93.
- Angwin J., Larson J., Mattu S., Kirchner L. (2016). Machine Bias: There's Software Used Across the Country to Predict Future Criminals. And It's Biased Against Blacks // *ProPublica*. May 23. <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing> (access date: 17.03.2026).
- Anzer G., Bauer P. (2021). A Goal Scoring Probability Model for Shots Based on Synchronized Positional and Event Data in Football (Soccer) // *Frontiers in Sports and Active Living*. Vol. 3. Art. 624475. DOI: 10.3389/fspor.2021.624475.
- Ariai F., Mackenzie J., Demartini G. (2025). Natural Language Processing for the Legal Domain: A Survey of Tasks, Datasets, Models, and Challenges // *ACM Computing Surveys*. Vol. 58. No. 6. Art. 163. Pp. 1–37. DOI: 10.1145/3777009.
- Baumer B.S., Jensen S.T., Matthews G.J. (2015). OpenWAR: An open source system for evaluating overall player performance in Major League Baseball // *Journal of Quantitative Analysis in Sports*. Vol. 11. No. 2. Pp. 69–84. DOI: 10.1515/jqas-2014-0098.
- Bradley R.A., Terry M.E. (1952). Rank Analysis of Incomplete Block Designs: The Method of Paired Comparisons // *Biometrika*. Vol. 39. No. 3–4. Pp. 324–345. DOI: 10.1093/biomet/39.3-4.324.
- Brave S.A., Butters R.A., Roberts K.A. (2019). Uncovering the sources of team synergy: Player complementarities in the production of wins // *Journal of Sports Analytics*. Vol. 5. No. 4. Pp. 247–279. DOI: 10.3233/JSA-190248.
- Cattelan M. (2012). Models for Paired Comparison Data: A Review with Emphasis on Dependent Data // *Statistical Science*. Vol. 27. No. 3. Pp. 412–433. DOI: 10.1214/12-STS396.
- Chen L.M., Staiger D.O., Birkmeyer J.D., Ryan A.M., Zhang W., Dimick J.B. (2013). Composite Quality Measures for Common Inpatient Medical Conditions // *Medical Care*. Vol. 51. No. 9. Pp. 832–837. DOI: 10.1097/MLR.0b013e31829fa92a.
- Chesterman S. (2020). Do Better Lawyers Win More Often? Measures of Advocate Quality and Their Impact in Singapore's Supreme Court // *Asian Journal of Comparative Law*. Vol. 15. No. 2. Pp. 250–280. DOI: 10.1017/asjcl.2020.13.
- Chetty R., Friedman J.N., Rockoff J.E. (2014a). Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates // *American Economic Review*. Vol. 104. No. 9. Pp. 2593–2632. DOI: 10.1257/aer.104.9.2593.
- Chetty R., Friedman J.N., Rockoff J.E. (2014b). Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood // *American Economic Review*. Vol. 104. No. 9. Pp. 2633–2679. DOI: 10.1257/aer.104.9.2633.
- Dressel J., Farid H. (2018). The accuracy, fairness, and limits of predicting recidivism // *Science Advances*. Vol. 4. No. 1. Art. eaao5580. DOI: 10.1126/sciadv.aao5580.
- Efron B., Morris C. (1975). Data Analysis Using Stein's Estimator and its Generalizations // *Journal of the American Statistical Association*. Vol. 70. No. 350. Pp. 311–319. DOI: 10.2307/2285814.
- Galanter M. (1974). Why the «Haves» Come Out Ahead: Speculations on the Limits of Legal Change // *Law & Society Review*. Vol. 9. No. 1. Pp. 95–160. DOI: 10.2307/3053023.
- Ghimire S., Ehrlich J.A., Sanders S.D. (2020). Measuring individual worker output in a complementary team setting: Does regularized adjusted plus minus isolate individual NBA player contributions? // *PLOS ONE*. Vol. 15. No. 8. Art. e0237920. DOI: 10.1371/journal.pone.0237920.
- Grassetti L., Bellio R., Di Gaspero L., Fonseca G., Vidoni P. (2021). An extended regularized adjusted plus-minus analysis for lineup management in basketball using play-by-play data // *IMA Journal of Management Mathematics*. Vol. 32. No. 4. Pp. 385–409. DOI: 10.1093/imaman/dpaa022.
- Greiner D.J., Pattanayak C.W. (2012). Randomized Evaluation in Legal Assistance: What Difference Does Representation (Offer and Actual Use) Make? // *Yale Law Journal*². Vol. 121. No. 8. Pp. 2118–2208.
- Hall B.L., Huffman K.M., Hamilton B.H., Paruch J.L., Zhou L., Richards K.E., Cohen M.E., Ko C.Y. (2015). Profiling Individual Surgeon Performance Using Information from a High-Quality Clinical Registry: Opportunities and Limitations // *Journal of the American College of Surgeons*. Vol. 221. No. 5. Pp. 901–913. DOI: 10.1016/j.jamcollsurg.2015.07.454.
- Herbrich R., Minka T., Graepel T. (2006). TrueSkill™: A Bayesian Skill Rating System // *Advances in Neural Information Processing Systems*. Vol. 19. Pp. 569–576. DOI: 10.5555/2976456.2976528.
- Karimi-Haghighi M., Castillo C. (2021). Enhancing a recidivism prediction tool with machine learning: effectiveness and algorithmic fairness // *Proceedings of the 18th International Conference on Artificial Intelligence and Law (ICAAIL '21)*. Pp. 210–214. DOI: 10.1145/3462757.3466150.
- Katz D.M., Bommarito M.J. II, Blackman J. (2017). A general approach for predicting the behavior of the Supreme Court of the United States // *PLOS ONE*. Vol. 12. No. 4. Art. e0174698. DOI: 10.1371/journal.pone.0174698.

² Йельский университет (Yale University) (США) признан нежелательной организацией в России.

- Kazun A.P. (2024). Can Artificial Intelligence Predict Court Decisions? A Systematic Review of International Research // *Monitoring obshchestvennogo mneniya: ekonomicheskie i sotsial'nye peremeny*. No. 5. Pp. 100-122. DOI: 10.14515/monitoring.2024.5.2580. (In Russ.).
- Koren Y., Bell R., Volinsky C. (2009). Matrix Factorization Techniques for Recommender Systems // *Computer*. Vol. 42. No. 8. Pp. 30–37. DOI: 10.1109/MC.2009.263.
- Krell R.W., Hozain A., Kao L.S., Dimick J.B. (2014). Reliability of Risk-adjusted Outcomes for Profiling Hospital Surgical Quality // *JAMA Surgery*. Vol. 149. No. 5. Pp. 467–474. DOI: 10.1001/jamasurg.2013.4249.
- Lazarus R.J. (2008). Advocacy Matters Before and Within the Supreme Court: Transforming the Court by Transforming the Bar // *Georgetown Law Journal*. Vol. 96. No. 5. Pp. 1487–1564.
- Macdonald B. (2012). Adjusted Plus-Minus for NHL Players Using Ridge Regression with Goals, Shots, Fenwick, and Corsi // *Journal of Quantitative Analysis in Sports*. Vol. 8. No. 3. Pp. 1–24. DOI: 10.1515/1559-0410.1447.
- Maruthappu M., Carty M.J., Lipsitz S.R., Wright J., Orgill D., Duclos A. (2014). Patient- and surgeon-adjusted control charts for monitoring performance // *BMJ Open*. Vol. 4. No. 1. Art. e004046. DOI: 10.1136/bmjopen-2013-004046.
- Menezes-Neto E.J., Clementino M.B.M. (2022). Using deep learning to predict outcomes of legal appeals better than human experts: A study with data from Brazilian federal courts // *PLOS ONE*. Vol. 17. No. 7. Art. e0272287. DOI: 10.1371/journal.pone.0272287.
- Minyust Rossii. (2026). *Unified State Register of Advocates*. URL: <https://minjust.gov.ru/ru/pages/advokaty-rossijskoj-federacii/> (accessed: 17.03.2026). (In Russ.).
- Mojon A., Mahari R., Lera S.C. (2025). Data-driven law firm rankings to reduce information asymmetry in legal disputes // *Nature Computational Science*. Vol. 5. No. 11. Pp. 1010–1016. DOI: 10.1038/s43588-025-00899-2.
- Nelson M.J., Epstein L. (2022). Human Capital in Court: The Role of Attorney Experience in US Supreme Court Litigation // *Journal of Law and Courts*. Vol. 10. No. 1. Pp. 61–85. DOI: 10.1086/714577.
- Paneyakh E.L. (2012). Practical logic of judicial decision-making: discretion under pressure and compromises at the expense of the defendant // *How judges make decisions: empirical studies of law / edited by V.V. Volkov*. — M.: Statut. Pp. 107–127. (In Russ.).
- Pitches D.W., Mohammed M.A., Lilford R.J. (2007). What is the empirical evidence that hospitals with higher risk-adjusted mortality rates provide poorer quality care? A systematic review of the literature // *BMC Health Services Research*. Vol. 7. Art. 91. DOI: 10.1186/1472-6963-7-91.
- Rosili N.A.K., Hassan R., Zakaria N.H., Kasim S., Che Rose F.Z., Sutikno T. (2021). A systematic literature review of machine learning methods in predicting court decisions // *IAES International Journal of Artificial Intelligence*. Vol. 10. No. 4. Pp. 1091–1102. DOI: 10.11591/ijai.v10.i4.pp1091-1102.
- Saveliev D., Kuchakov R. (2024). The Russian Legislative Corpus // *arXiv preprint arXiv:2406.04855*. 2024. DOI: 10.48550/arXiv.2406.04855.
- Titaev K. (2017). Pretrial detention in Russian criminal courts: a statistical analysis // *International Journal of Comparative and Applied Criminal Justice*. Vol. 41. No. 3. Pp. 145–161. DOI: 10.1080/01924036.2016.1239117.
- Tollenaar N., van der Heijden P.G.M. (2019). Optimizing predictive performance of criminal recidivism models using registration data with binary and survival outcomes // *PLOS ONE*. Vol. 14. No. 3. Art. e0213245. DOI: 10.1371/journal.pone.0213245.
- Travaini G.V., Pacchioni F., Bellumore S., Bosia M., De Micco F. (2022). Machine Learning and Criminal Justice: A Systematic Review of Advanced Methodology for Recidivism Risk Prediction // *International Journal of Environmental Research and Public Health*. Vol. 19. No. 17. Art. 10594. DOI: 10.3390/ijerph191710594.
- Vasconcelos C.C., Oliveira E.W., Netto W.L. (2018). The Impact of Attorneys on Judicial Decisions: Empirical Evidence from Civil Cases // *International Journal for Court Administration*. Vol. 9. No. 2. Pp. 32–42. DOI: 10.18352/IJCA.244.
- Vinogradov A.S. (2025). Accusatory Bias in Preliminary Criminal Proceedings: Causes and Ways to Overcome // *Zhurnal pravovykh i ekonomicheskikh issledovaniy (Vestnik GIEF)*. Vol. 59. No. 1. Pp. 133–138. DOI: 10.26163/gief.2025.59.52.016. (In Russ.).
- Volkov V.V. (2014). *Socioeconomic status and differences in sentencing: data on the consideration of criminal cases in the courts of the Russian Federation* / Preprint of the Institute of Law Enforcement Problems at the European University at St. Petersburg. — SPb.: IPP EUSP (In Russ.).
- Volkov V. (2021). The prosecutor effect in trials for petty violent offences in Russia // *International Journal of Comparative and Applied Criminal Justice*. Vol. 45. No. 2. Pp. 207–220. DOI: 10.1080/01924036.2020.1732434.
- Zhuchkova S.V., Devyatnikov V.Yu., Kazun A.P., Belov M.D., Sidorova O.I. (2025). Texts of Court Verdicts as a Data Source for Empirical Legal Studies in Russia // *Monitoring obshchestvennogo mneniya: ekonomicheskie i sotsial'nye peremeny*. No. 2. Pp. 170–192. DOI: 10.14515/monitoring.2025.2.2575. (In Russ.).
- Zhuchkova S., Kazun A. (2023). Exploring Gender Bias in Homicide Sentencing: An Empirical Study of Russian Court Decisions Using Text Mining // *Homicide Studies*. Advance online publication. DOI: 10.1177/10887679231217159.

Казун Антон Павлович

ORCID: 0000-0002-0091-5388

akazun@hse.ru

Anton Kazun

Cand. Sci. (Soc.), Director, Institute for Industrial and Market Studies, HSE University (Moscow, Russia)

ORCID: 0000-0002-0091-5388

akazun@hse.ru

Девятников Вадим Юрьевич

ORCID: 0009-0009-5278-7232

vdevyatnikov@hse.ru

Vadim Devyatnikov

Research Assistant, International Center for the Study of Institutions and Development, HSE University (Moscow, Russia)

ORCID: 0009-0009-5278-7232

vdevyatnikov@hse.ru

Белов Михаил Дмитриевич

ORCID: 0000-0001-8848-2726

mikaelbelov@gmail.com

Mikhail Belov

PhD student, University of California, Los Angeles, Los Angeles, USA

ORCID: 0000-0001-8848-2726

mikaelbelov@gmail.com

INFORMATION ASYMMETRY IN THE LEGAL SERVICES MARKET: CREATING A RANKING USING MACHINE LEARNING METHODS³

Abstract. How can you tell whether a defense attorney is any good? In this article, we describe a methodology for measuring the quality of legal representation by comparing machine learning predictions based on objective case characteristics with actual case outcomes. Our premise is that if an attorney's cases consistently end better than expected, they are probably doing their job well. Similar logic has long been used in other fields: it is how teachers are evaluated in the economics of education, surgeons in healthcare, and players in professional sports. We adapt these approaches to criminal proceedings. Such a predictive model can be trained on Russian court data from 2010–2025 and would account for the fact that a criminal case passes through several sequential procedural decision points. The resulting rating is adjusted for teamwork among multiple defense attorneys, the stage of proceedings, and small caseloads. We discuss the weaknesses of the approach: most importantly, the pretrial stage is invisible to us, published decisions are incomplete, and skilled attorneys may systematically select more difficult cases. Nevertheless, even in this form, such an evaluation tool can provide clients with a meaningful and objective benchmark when choosing a defense attorney, which could substantially improve the quality of information in the legal services market.

Keywords: *defense attorney performance evaluation; judicial decision prediction; machine learning; information asymmetry; criminal defense.*

JEL: K14, K41, C53, C55, D82.

³ The study was carried out within the framework of the Basic Research Program of HSE University, No. HSE-BR-2025-034.