STRATEGIC MANAGEMENT PERSPECTIVE ON LITIGATION PREDICTION FOR US BODILY-INJURY INSURANCE CLAIMS

Major Project Report Submitted
in Partial Fulfillment of the Requirements for the
Degree of

MASTER OF BUSINESS ADMINISTRATION (EXECUTIVE)

in

DATA SCIENCE AND ANALYTICS

by

NEERAJ SHARMA
(Roll No. 24/UEMBA/07)

Under the Supervision of
Prof. MANOJ KUMAR SHARMA

UNIVERSITY SCHOOL OF MANAGEMENT & ENTREPRENEURSHIP
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Main Bawana Road, Delhi-110042, India

July, 2026

ACKNOWLEDGEMENT

I would like to take this opportunity to express my heartfelt gratitude to everyone who has contributed to the successful completion of this project, “Strategic Management Perspective on Litigation Prediction for US Bodily-Injury Insurance Claims”.

First and foremost, I extend my deepest gratitude to my project guide, Professor Manoj Kumar Sharma, for his invaluable guidance, insightful feedback, and constant encouragement throughout the course of this research. His expertise and mentorship have been instrumental in shaping the direction and quality of this study.

I am also immensely thankful to the University School of Management & Entrepreneurship, Delhi Technological University and the faculty of the institute for providing me with the resources, knowledge, and support needed to undertake this project. The academic environment and infrastructure at the institution have been crucial in facilitating my research.

Additionally, I am grateful to my peers and colleagues for their constructive discussions and moral support during the challenging phases of the project. Their camaraderie has been a source of motivation throughout this journey.

Finally, I owe a great deal of gratitude to my family for their unwavering support, patience and encouragement during this endeavour. Their belief in me has been a constant source of strength.

This project would not have been possible without the collective efforts and contributions of all these individuals and institutions. To all who have supported me directly or indirectly, I extend my sincere thanks.

Name: Neeraj Sharma
Roll No: 24/UEMBA/07

DECLARATION

I Neeraj Sharma hereby certify that the work which is being presented in the capstone report entitled “Strategic Management Perspective on Litigation Prediction for US Bodily-Injury Insurance Claims” in partial fulfillment of the requirements for the award of the degree of Masters of Business Administration (Executive) in Data Science and Analytics, submitted in the University School of Management & Entrepreneurship, Delhi Technological University is an authentic record of my own work carried out during the period from June’2026 to July’2026 under the supervision of Prof. Manoj Kumar Sharma.

The matter presented in the project report has not been submitted by me for award of any other degree of this or any other institute.

Name: Neeraj Sharma
Roll No: 24/UEMBA/07

CERTIFICATE BY THE SUPERVISOR

Certified that Neeraj Sharma (Roll No. 24/UEMBA/07) has carried out their capstone work presented in this project report titled “Strategic Management Perspective on Litigation Prediction for US Bodily-Injury Insurance Claims” for the award of Masters of Business Administration (Executive) Data Science and Analytics from University School of Management and Entrepreneurship, Delhi Technological University, under my supervision. The capstone embodies results of original work, and studies are carried out by the student himself and the contents of the project report do not form the basis for the award of any other degree to the candidate or to anybody else from this or any other University/Institution.

Prof. Manoj Kumar Sharma
University School of Management & Entrepreneurship,
Delhi Technological University

ABSTRACT

Litigated bodily-injury (BI) claims are among the costliest events in US property-casualty insurance: claims involving attorneys settle for a multiple of comparable unrepresented claims, close more slowly, and absorb disproportionate loss adjustment expense, a dynamic intensified by the industry-wide trend of "social inflation." Because attorney-retention decisions are made in the first days after a loss, and because early, fair claim handling measurably reduces the probability of representation, the ability to predict litigation at first notice of loss (FNOL) is not merely an analytics capability — it is a strategic one. This study examines that capability through an explicitly strategic-management lens, asking not only whether litigation can be predicted, but how the prediction reconfigures an insurer's value chain to produce durable cost and differentiation advantage.

The project is executed under the CRISP-DM methodology. In the business-understanding phase, Porter's value-chain framework is used to map litigation prediction onto every primary and support activity of the claim's operation, from FNOL intake through settlement, pricing feedback, counsel procurement, and governance. In the data, modelling, and evaluation phases, a 12,000-claim synthetic dataset — schema-faithful to public US closed-claim databases (Texas Department of Insurance; Florida Office of Insurance Regulation) — is used to train and compare a regularized logistic regression against a calibrated gradient-boosting classifier under a leakage-safe chronological validation design. The selected model attains a held-out ROC-AUC of 0.79 and a top-decile lift of 3.42×: the highest-scored 10% of claims contains roughly one-third of all future litigation. Every score is accompanied by a calibrated risk band with a recommended handling action and a counterfactual, adjuster-actionable explanation. In the deployment phase, the capability is delivered through a Streamlit application and an agentic natural-language interface (LangChain with Anthropic's Claude) in which a tool-calling agent plans multi-step analyses over five deterministic, auditable tools and is architecturally prevented from inventing numbers.

The study concludes that the strategic value of litigation prediction lies less in the algorithm than in the linkages it activates across the value chain — triage, staffing, settlement timing, counsel procurement, reserving, and pricing — which simultaneously widen margin and grow market share while remaining difficult for competitors to imitate. Ethical and regulatory guardrails, aligned with the NAIC's model bulletin on insurers' use of artificial intelligence, are treated as constitutive of the strategy rather than constraints upon it.

TABLE OF CONTENTS

ACKNOWLEDGEMENT ii
DECLARATION iii
CERTIFICATE BY THE SUPERVISOR iv
ABSTRACT v
TABLE OF CONTENTS vi
LIST OF TABLES vii
LIST OF FIGURES viii
LIST OF ABBREVIATIONS ix
CHAPTER 1: INTRODUCTION 1
CHAPTER 2: OBJECTIVE OF STUDY 3
CHAPTER 3: LITERATURE REVIEW 5
CHAPTER 4: RESEARCH METHODOLOGY 9
CHAPTER 5: RESULTS 18
CHAPTER 6: CONCLUSIONS 23
APPENDICES 26
REFERENCES 32
PLAGIARISM VERIFICATION 34

LIST OF TABLES

Table 4.1: CRISP-DM phases mapped to project activities 11
Table 4.2: Feature inventory (21 predictors and the target) 14
Table 4.3: Risk bands and recommended handling actions 17
Table 4.4: Agent tools and their grounding 19
Table 5.1: Held-out performance (chronological test set, 2,400 claims, base rate 16.2%) 20

LIST OF FIGURES

Figure 1.1: Lifecycle of a US bodily-injury insurance claim..... 1
Figure 4.1: Porter value chain mapped to litigation prediction (Exhibit, landscape) 12
Figure 4.2: Litigation rate by state versus book average (12,000 synthetic claims). 15
Figure 4.3: Litigation rate by attorney involvement at FNOL and liability posture. 15
Figure 4.4: Medical specials (log scale) by litigation outcome: shifted but heavily overlapping distributions..... 16
Figure 4.5: System architecture: both user-facing surfaces call the same scoring core... 18
Figure 5.1: ROC curves for both models on the held-out test set. 20
Figure 5.2: Reliability (calibration) curve, quantile-binned..... 21
Figure 5.3: Actual litigation rate by model-score decile; top-decile lift 3.42×..... 22
Figure 5.4: Predicted-probability distributions by outcome, with risk-band thresholds.. 22
Figure 5.5: Permutation feature importance on the held-out set (top 12). 23
Figure 5.6: Counterfactual decomposition of the worked example (predicted 76.6%, Critical band)..... 23

LIST OF ABBREVIATIONS

AI	Artificial Intelligence
ALAE	Allocated Loss Adjustment Expense
API	Application Programming Interface
AUC	Area Under the Curve
BI	Bodily Injury
CAS	Casualty Actuarial Society
CRISP-DM	Cross-Industry Standard Process for Data Mining
EDA	Exploratory Data Analysis
FNOL	First Notice of Loss
GBM	Gradient Boosting Machine
III	Insurance Information Institute
IRC	Insurance Research Council
LAE	Loss Adjustment Expense
LLM	Large Language Model
ML	Machine Learning
NAIC	National Association of Insurance Commissioners
NLP	Natural Language Processing
OIR	(Florida) Office of Insurance Regulation
PR-AUC	Precision-Recall Area Under the Curve
ROC	Receiver Operating Characteristic
SHAP	SHapley Additive exPlanations
TDI	Texas Department of Insurance
ULAE	Unallocated Loss Adjustment Expense
XAI	Explainable Artificial Intelligence

CHAPTER 1: INTRODUCTION

1.1 Industry context: social inflation and the litigation problem

US property-casualty insurers face a sustained rise in liability claim costs that outpaces economic inflation — a phenomenon the industry calls social inflation. Its drivers are well documented: growing attorney involvement in injury claims, expanding attorney advertising, third-party litigation funding, and a jury environment producing more very large ("nuclear") verdicts. the real life example of this trend of litigation. If a bodily injury claimant hires an attorney, the "economic" aspects of the claim undergo an essentially radical shift the data universally indicate that settled claims handled by attorneys are paid out many times greater than similarly claims filed by the plaintiffs alone, take much, much longer to settle and require substantially higher loss adjustment expense. Litigation is one of the largest controllable costs in the casualty lines: controllable, since litigation is not solely exogenous early, empathetic contact with claimant, prompt liability determination and reliable timely evaluation of a settlement have a statistically significant impact on preventing plaintiff counsel from entering a case.

1.2 The claim lifecycle and the intervention window

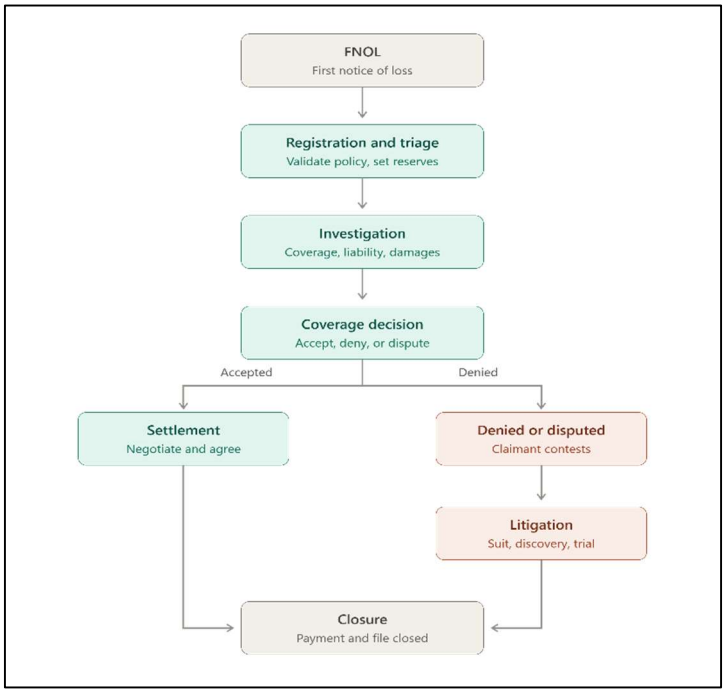


Figure 1.1: Lifecycle of a US bodily-injury insurance claim, from FNOL to closure.

Figure 1.1 traces the lifecycle of a US bodily-injury claim from first notice of loss (FNOL) through registration and triage, investigation, and the coverage decision, after which the claim follows one of two paths: negotiated settlement, or the denied/disputed branch that leads into litigation — suit, discovery, and trial — before eventual closure. The strategic significance of the diagram lies in its timing. The fork between the settlement path and the litigation path is substantially determined in the earliest stages, when the claimant is deciding whether to seek representation. A prediction that arrives at FNOL — while the claim is still in triage — can therefore change which branch the claim takes; a prediction that arrives after the fork can only describe it. This is why the system studied here scores claims at intake rather than at any later point.

1.3 Why a strategic management perspective?

Most models dealing with claim scoring focus solely on the prediction aspects. This paper contends, however, that this is where most authors stop: "prediction is just the beginning." An algorithm itself brings no monetary benefit. To generate profit, downstream activities such as staffing, data quality, and timely disposition must be rearranged in response to the score, so intake obtains good, usable data; triage sends to the correct channel on the basis of score; staffing provides the trained staff at the desk the scores promise; settlement procedures give them the authority to make timely, fair offers; counsel is employed with knowledge of the litigation risk; reserving and pricing reflect the score; and governance keeps the entire process aligned with regulatory and ethical standards.

1.4 Problem statement

Given information plausibly available to the claim handler at or shortly after the first notification of loss, a machine will predict, explain in a format that would allow the claims handler to follow the reasoning with the ability to accept or reject the rationale, and predict the probability that a U.S bodily injury claim will become litigated, and Porter chain analysis of the ROI of this prediction.

1.5 Scope and organization of the report

Chapter Two explains the purpose of this case study. Chapter Three reports a selective literature review of the three distinct fields: strategic management, insurance economics and AI/machine learning. Chapter Four describes the data creation and the research methodology. Data consists of fully synthetic records following all structural constraints of the industry schema. Chapters Five and Six show the results obtained from the data analysis via AI/machine learning algorithms, value chains and analysis, limitations and future directions, and finally some ethical concerns that should be paid attention to.

CHAPTER 2: OBJECTIVE OF STUDY

2.1 Primary Objective

An end-to-end litigation prediction tool for US bodily injury insurance claims needs to be created, built and tested, along with an analysis, from a value chain, strategic management perspective on how such an ability creates sustainable competitive advantage in profits and market share.

2.2 Specific objectives

SO1 (Predictive): Train classifiers and compare them on FNOL-feasible features, in a leak-safe chronological validation protocol aimed at producing material lift for concentration of litigation risk in the top decile (geq.3).

SO2 (Probabilistic): Generate well calibrated probabilities for threshold based risk bands suitable for operational application, measured using Brier score and reliability curves.

SO3 (Explainable): Produce a counterfactual explanation that an adjuster can act upon along with every score, and validate global determinants against the claims literature.

SO4 (Agentic): Develop a tool calling language model agent that can answer questions at a portfolio or claim level via only grounded and auditable tool calls.

SO5 (Deliverable): Build a Streamlit app to offer a monitor portfolio, scoring on an individual claim, triage batch and conversational analysis.

SO6 (Strategic): Overlain on Porter's value chain of both main and support activities and describe how the capability would lead to profit and market share benefits and how the benefits could be defended.

SO7 (Responsible): Incorporate fairness, ethical, and regulatory (conforming to the NAIC bulletin on AI) aspects into both system design and recommendations from strategy.

2.3 Research questions

RQ1: What is the FNOL prediction accuracy for litigation given the information plausibly available at that time?

RQ2: To what extent do claim-, behavioral-, and legal-environment related variables contribute to litigation risk, and do they align with established empirical literature?

RQ3: Can trust be engendered in model predictions to set stable operational thresholds?

RQ4: In which dashboards and agentic interfaces should the capability be applied so as to promote use, auditable and safety.

RQ5: Through what activities and value chain relationships does the capability produce benefits, and why are those difficult to imitate?

CHAPTER 3: LITERATURE REVIEW

The review proceeds in nine strands, moving from the strategic-management foundations that frame the study, through the insurance and legal-empirical literatures that define the problem, to the machine-learning, governance, and methodological literatures that equip the solution. Each strand closes by stating what the present study takes from it; the chapter ends with a synthesis and gap statement.

3.1 Strategic management foundations

Porter (1985) introduced the value chain as the instrument for disaggregating a firm into strategically relevant activities — five primary, four support — whose individual cost behaviour and interdependencies ("linkages") determine competitive advantage. Two of Porter's arguments carry particular weight here. First, advantage arises not only within activities but from the coordination between them, and linkages are harder for rivals to observe and replicate than any single activity. Second, technology development is a support activity that can transform the entire chain when it changes how primary activities are performed. Porter (1996) sharpened the point: operational effectiveness alone is imitable, and durable strategy consists in a system of mutually reinforcing activities. Barney's (1991) resource-based view supplies the complementary test — resources sustain advantage when they are valuable, rare, imperfectly imitable, and non-substitutable — a test that a coordinated model-plus-workflow system can pass where a stand-alone algorithm cannot. From this strand the study takes its central analytical claim: the litigation model must be analyzed as a value-chain reconfiguration, not as an isolated artifact.

3.2 Attorney involvement and the economics of injury claims

Labelled as injury claim resolution the empirical subject was first introduced via the RAND Institute for Civil Justices extensive evaluations of tort damages in the US Hensler et al. 1991. In the context of this type of research, the recurring Attorney Involvement in Auto Injury Claims studies by the Insurance Research Council offered the greatest direct precedent for this project because over multiple years and closed claim surveys they track continuously increasing attorney representation rates in US BI claims, while still witnessing a sustained outcome effect- Attorney represented claims are resolved in longer time spans, incurring larger total costs, although to a degree offset by reduced contingency fee paid out to the claimant (Insurance Research Council, various years). These two results lead to clear design specifications: attorney involvement should be the target for predicting and any useful predictor needs to be applicable at the FNOL where decisions regarding claimant representation cannot be based upon intake- era variables.

3.3 Social inflation and the legal environment

The Insurance Information Institute, jointly with the Casualty Actuarial Society, has traced social inflation to growing attorney involvement, attorney advertising, third-party litigation funding, and shifting jury attitudes, and has shown measurable acceleration in casualty loss-development patterns (Lynch & Moore, 2022). Swiss Re Institute's sigma research reaches parallel conclusions and emphasizes venue concentration: a small set of jurisdictions accounts for a disproportionate share of outsized verdicts (Swiss Re Institute, 2024). This strand contributes the study's legal-environment features — state, urban venue, disputed liability, demand-to-limit ratio — and later supplies the face-validity benchmark against which learned feature importances are checked.

3.4 Empirical legal studies on closed-claim databases

Because insurer claim files are confidential, state closed-claim databases — principally the Texas Department of Insurance closed-claim reports and the Florida Office of Insurance Regulation's collections — are the main public windows into US claim outcomes, and they support a substantial peer-reviewed literature. Most prominently, Black, Silver, Hyman and Sage (2005) used TDI closed-claim data to analyze medical-malpractice outcomes over 1988–2002, finding stable rather than crisis-level dynamics — work that influenced national tort-reform debates and demonstrated that closed-claim data can sustain rigorous empirical analysis. For the present study the strand plays a legitimizing role: the synthetic dataset deliberately mirrors these databases' schemas, so the demonstrated pipeline can be re-pointed at data on which real scholarship already runs.

3.5 Predictive modelling in insurance claims

Machine learning entered claims operations primarily through fraud detection. Derrig (2002) framed insurance fraud as a statistical decision problem; Viaene et al. (2002), benchmarking classifiers on automobile injury-claim fraud, found — notably for this study — that logistic regression remained competitive with more complex learners on tabular claim data. The modern actuarial toolkit is systematized in Wüthrich and Buser's widely used lecture notes on data analytics for non-life insurance, with gradient-boosted trees (Friedman, 2001) the de facto standard for tabular problems. Published litigation models specifically are scarce — practitioner white papers describe attorney-involvement scoring but disclose neither data nor validated performance — which forms part of the gap this study addresses. The applied literature also supplies a methodological warning central to this study's design: Kaufman et al. (2012) formalize leakage — inclusion of information unavailable at prediction time — as among the most

common and damaging errors in applied data mining, motivating both the chronological train/test split and the restriction to FNOL-plausible features.

3.6 Calibration of Probabilities

A triage system that maps probabilities to action thresholds requires reliable estimates of probabilities, and this cannot be taken for granted from the literature on calibration. Platt (1999) introduced parametric (sigmoid) rescaling of classifier outputs; Zadrozny and Elkan (2002) provided the non-parametric isotonic method of Platt in Platt that is the basis for this analysis; Niculescu-Mizil and Caruana (2005) have also empirically shown that boosted outputs tend to be mis-calibrated and that this can be effectively remedied by calibration. Analysis follows this line of research: the boosting contestant was isotonically calibrated and all quality metrics, the Brier score, and reliability plots for all ranks are considered, in addition to ranking metrics.

3.7 Explainable AI

In this sense permutation importance (the decrease in predictive performance when a feature is shuffled) is a method that was devised by Breiman (2001) and is the only truly model-agnostic technique in the list of those we consider. Within this local approach, there are three dominant types of attribution: additive model-agnostic explanations that rest on game-theoretic definitions (SHAP; Lundberg and Lee, 2017), local linear surrogates (LIME; Ribeiro, Singh & Guestrin, 2016) and local counterfactual explanations first proposed in algorithmic accounting in Wachter, Mittelstadt, and Russell (2017), which aim to explain the prediction by describing how it changes under differences from reality, a body of work codified in Molnar (2022). The study's swap-to-baseline strategy constitutes a simplified approach from within the counterfactual family -– it eschews coalition attribution from axiomatic SHAP attribution for no extra complexity, perfect model faithfulness and an output framed directly in claim-handling terms. For clarity of grounds for choice, we articulate both.

3.8 AI governance in insurance and tool-using language model agents

The move to moving from principles to expectations is already underway by US regulators; NAIC's Bulletin On the Use of Artificial Intelligence by Insurers (2023), largely followed by a majority of states, sets forth AI governance programs along with testing and records to be documented, explainability with market conduct examinations for consumer outcomes. Within actuarial science, discrimination-free prediction has been axiomatized by Lindholm, Richman, Tsanakas, Wüthrich (2022), offering precise terms for pre-production fairness tests required, along with others. On the agentic dimension, Yao et al. (2023) prototyped ReAct by alternating reason steps and action steps with tools, establishing that grounding cuts down hallucination, while Schick et al.

(2023) leveraged the architecture of the Toolformer to show that tools can be learned to invoke when necessary. The agent used in the study follows this path exactly, with one additional assumption: every claim made must be expressed as tool output.

3.9 Using CRISP-DM as the Methodological Framework

Developed by an industry consortium a six phase iterative process model, the Cross-Industry Standard Process for Data Mining (CRISP-DM) specifically outlined business understanding, data understanding, data preparation, modeling, evaluation and deployment (Chapman et al., 2000; Wirth & Hipp, 2000) in order to keep data science projects focused on a business purpose. Twenty years of use has proven it to be the established 'way of doing things' for real-world data science projects, with recent summaries noting its resilience but highlighting the imperative to strengthen the deployment and governance phases in the context of machine-learning (Martínez-Plumed et al., 2021). CRISP-DM is perhaps more tailored than most methods to this case study given its initial phase of business understanding is where the Porter value chain analysis fits naturally; strategic diagnosis is not a separate section but the first concrete activity. As such, chapter 4 will proceed phase-by-phase. A methodological note for completeness is that CRISP-DM is a methodology for executing projects, rather than reviewing the literature it is, as stated above, cited as this study's methodology and used throughout Chapter 4.

3.10 Synthesis and research gap

Putting the pieces together, the strands of evidence establish that attorney engagement is the cost driver which US injury plaintiffs can most affect and that it's affected in the earliest stage of the case (3.2–3.3); that closed-claim data forms an effective empirical base (3.4); that each required technical instrument (tabular classifiers, leakage-protected testing and validation, calibration, explanation tools) is already available (3.5–3.7); that the expectations concerning governance and agentic behavior are already clearly described (3.8); and that the strategic management theory points towards sustainable competitive advantage emerging from integrated activity systems not isolated tools (3.1). The literature simply does not offer a public, open, reproducible, end-to-end litigation system considered as a strategic capability; the published, claim-level models of US injury litigation are proprietary and proprietary, and the available literature lacks a system calibrated for triage that provides per-claim explana–tions through conversation, documented in a governance structure, and analyzed with a clear value chain. This study fills that research gap by creating an open end-to-end litigation system on synthetic but schema faithful data, and providing the complete system as a reproducible baseline.

CHAPTER 4: RESEARCH METHODOLOGY

Chapman et al. (2000) broke down the process into six steps, yet it moves forward following a method plan laid out by CRISP-DM. Even after multiple interactions, the sequence shifts only slightly, still drawing its core from the initial model.

A quick reference showing how stage titles relate to real project tasks can be found in Table 4.1. Upon reflection later on, recurring patterns emerged - cycles appeared where outcomes influenced prior choices, such as how models were designed, highlighted clearly in Chapter 5. Despite these adjustments, the order stays true to where CRISP-DM originally placed each phase.

Table 4.1: CRISP-DM phases mapped to project activities

CRISP-DM phase	Realization in this study
1. Business understanding	Strategic diagnosis via Porter's value chain (Figure 4.1); litigation-cost problem definition; success criteria (Chapter 2)
2. Data understanding	Closed-claim schema study (TDI, FL OIR); synthetic data design; exploratory analysis (Figures 4.2–4.4)
3. Data preparation	Synthetic generation with documented drivers; leakage-safe preprocessing inside the model pipeline (§4.4)
4. Modelling	Regularized logistic regression baseline vs. calibrated histogram gradient boosting (§4.5)
5. Evaluation	Chronological hold-out; ROC-AUC, PR-AUC, Brier, decile lift; permutation importance; domain-logic tests (§4.6)
6. Deployment	Streamlit application; LangChain tool-calling agent; governance guardrails (§4.7, Figure 4.5, Table 4.4)

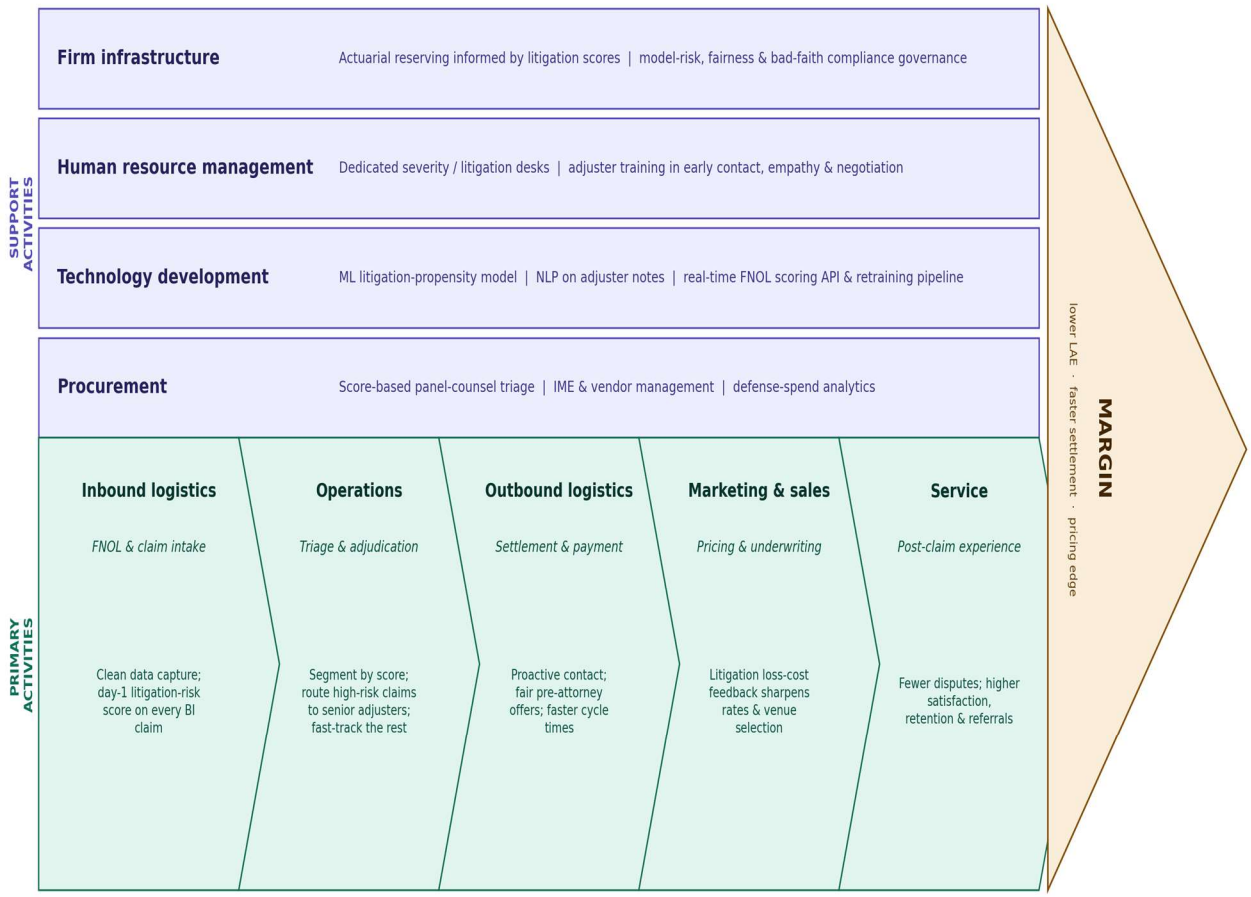
4.1 Research design overview

Construction followed by assessment shapes this research layout - a method where something new (prediction tools) gets built, then tested using saved data, while insights on real-world impact emerge through examination of how operations connect within a value chain. Instead of outside lab work, numbers come straight from tools the team ran themselves, relying on set number sequences and automated checks; judgments about relevance tie back to clear steps in commerce flow and relationships among them.

4.2 Phase 1 — Business understanding: the value-chain diagnosis

Starting from CRISP-DM's process, we begin with the business goal informing the analytical target. Here, the link appears clearly when applying Porter's value chain; the assertion functions much like a product moving through stages. From intake as uncensored data at first notice, it shifts form during handling and judgment processes. Outputs then emerge as payments settled and user engagement finished, while pricing decisions draw reference from outcomes returned after delivery. Feedback loops guide underwriting responses too. Though not generating output themselves, support roles - systems, people, tech creation, material purchase - enable continuous handling and maintain control. A visual representation showing this path appears in Figure 4.1, using traditional arrows linking step to step.

Figure 4.1: Porter value chain mapped to litigation prediction (US bodily-injury claims)



Reading: each activity shows its insurance-claims interpretation and the role litigation-propensity prediction plays in it. All linkages converge on margin.

The diagram reveals its strategy when unfolded step by step. Day one brings a risk score for each claim, sets the stage - attorneys choose involvement mostly right away after damage occurs. That score splits the workflow. Fewer complex cases slide into simple, hands-off systems or continue untouched through automated paths. Heavier risk cases turn instead to skilled reviewers, these handlers spread thin on volume, focused on reaching policyholders quickly, assessing claims reliably early - tackling expense from multiple sides at once. When payouts draw near, moderate settlements on difficult cases stop lawyers from stepping in, shortening overall processing duration. Understanding exactly where legal tendencies cluster becomes clearer through precise adjustments in insurance pricing. This clarity shapes how much risk a company will take on; what was once just cost savings now fuels expansion efforts. Handling claims thoughtfully from start helps reduce desire for legal representation. A calm, responsive approach builds trust, not only with customers but also within the system itself. Trust Deepens. People Stay. Word Spreads.

Accuracy in reserve calculations is highly dependent on systems below it infrastructure shapes the robustness of risk models at the desk while promoting fairness at the desktop through testing and paperwork. In addition, staff-focused efforts resulted in the formation of specialized teams that focus on severe cases and litigations; their performance incentives reward efficiency not delays in payment. New software tools also centralize logic, extract relevant information from unstructured adjuster comments via NLP, generate immediate risk scores, and monitor dynamic trends over time. Buying decisions apply the system's output - sorting law firms and suppliers by risk so financial data ties everything back home. Even if someone copies the code, matching real results isn't possible; replicating every part at once - clean data arrival, proper funneling, assigned workspaces, outside counsel plans, price feed loops, oversight structure - stands as a hurdle far tougher than duplication.

4.3 Phase 2 — Data understanding

4.3.1 Why synthetic data, and how it stays honest?

Insurer information per claim stays private; even though extracted closed claims appear in public, permission forms must be signed beyond any academic deadline. To move forward, authors produce synthetic data involving 12,000 personal injury cases governed by two rules. One rule: the structure copies actual forms found in Texas DOI and Florida OIR records, so each entry can draw directly from genuine releases and migration happens via corresponding columns only. From there, values emerge based on established drivers pulled from academic work reviewed earlier - lawyers often involved followed by issues around fault, hurt severity, location patterns, timing of reports and

care, pressures toward limits, signs of manipulation, protected statements, all mixed with strong randomness (standard deviation at 0.55). Real-world patterns appear with errors present; what is measured mirrors actual flaws in data. Starting litigation prevalence sits just under 16 percent through the intercept. No real-world sources were accessed; every number reported lives within this artificially built crowd.

4.3.2 Feature inventory

Table 4.2: Feature inventory (21 predictors and the target)

Group	Fields	Rationale
Claim & injury	claim_type, injury_type, injury_severity (1–5), claimant_age, claimant_gender	Stakes and context; med-mal and product lines litigate at structurally higher rates
Financial	medical_specials, lost_wages, initial_reserve, policy_limit, demand_to_limit_ratio	Economic stakes; demands near limits create time-limited-demand dynamics
Behavioural / process	report_lag_days, treatment_gap_days, prior_claims_count, attorney_at_fnol, recorded_statement_given, chiropractic_treatment, soft_fraud_indicators	Classic adjuster signals: late reporting, treatment gaps, build-up patterns, cooperation
Legal environment	state, urban_venue, disputed_liability	Venue litigation propensity and liability posture
Target	litigated (0/1)	Attorney-driven litigation; base rate ≈ 16%

4.3.3 Exploratory analysis

Three exploratory views validate face before modeling Three views, exploratory views, validate face before modeling Three exploratory views before modeling validate face Three views, exploratory views, validate face, before modeling Three exploratory views validate face before modelling Three exploratory views validate face before modeling Exploratory views validate face before modeling Three exploratory views establish face validity before modeling. 4.2 Litigation Rates by State: Georgia, Florida, and New York litigation rates are significantly higher than the book average while Indiana and North Carolina are below the book average, echoing the social inflation literature. 4.3 Interaction within the Domain: Attorney involvement at FNOL causes the litigation rate

to more than multiple, and this impact is magnified with a dispute as to liability. 4.4 Litigation claims in this sample have shifted toward more, and particularly "heavy", medical specials but there is great overlap in the medical specials as in the "specials", and the financial "stake" may not be sufficiently disparate to produce two distinct classes, which a multivariate model would establish.

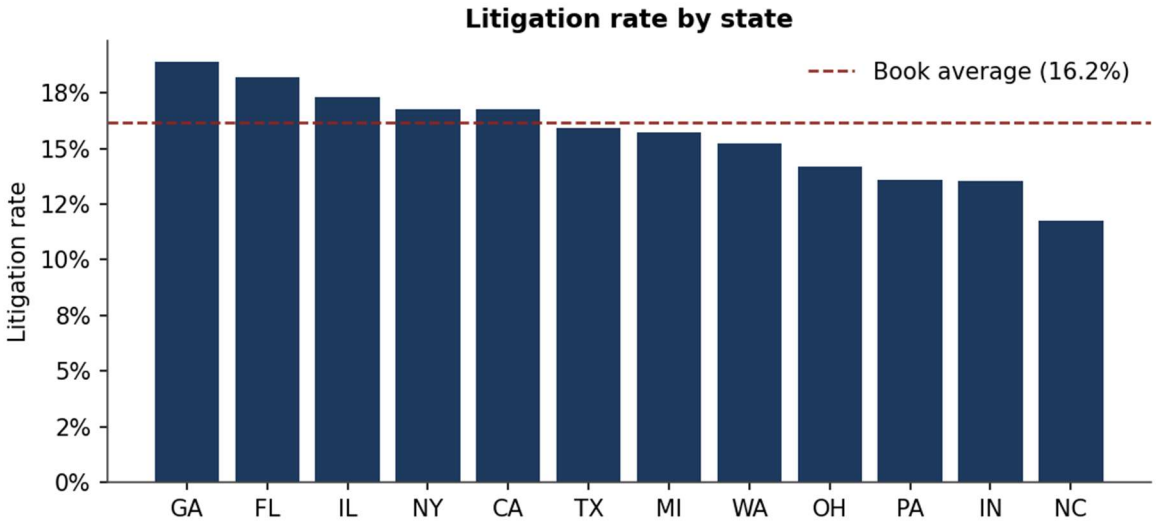


Figure 4.2: Litigation rate by state versus book average (12,000 synthetic claims).

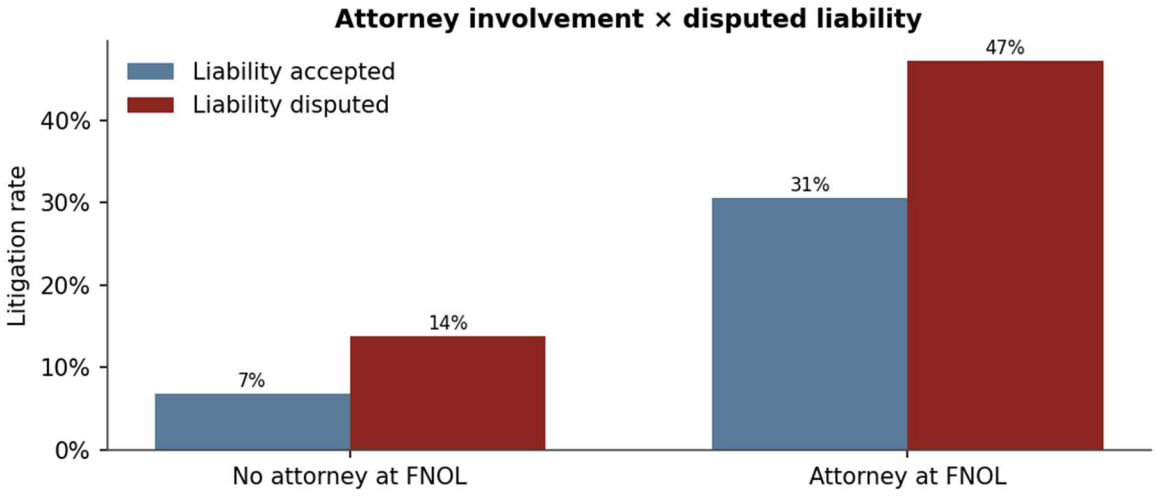


Figure 4.3: Litigation rate by attorney involvement at FNOL and liability posture.

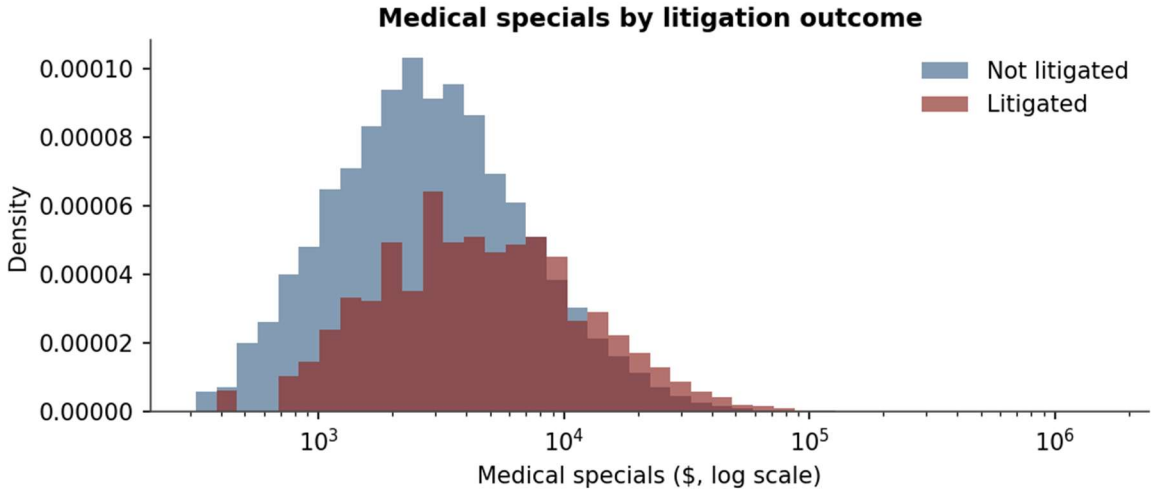


Figure 4.4: Medical specials (log scale) by litigation outcome: shifted but heavily overlapping distributions.

4.4 Phase 3 — Data preparation

Preprocessing is expressed as a scikit-learn ColumnTransformer and fitted inside the model pipeline, so scaling statistics and category vocabularies are learned from training data only and are persisted with the classifier as a single artifact. Numeric features are standardized; binary flags pass through; categoricals are one-hot encoded with unknown-safe handling so an unseen category degrades gracefully at scoring time. Fitting preprocessing inside the pipeline is the second of the study's two leakage protections (the first, the chronological split, is described under evaluation).

```
def build_preprocessor():
    return ColumnTransformer([
        ("num", StandardScaler(), NUMERIC_FEATURES),
        ("bin", "passthrough",   BINARY_FEATURES),
        ("cat", OneHotEncoder(handle_unknown="ignore"),
CATEGORICAL_FEATURES),
    ])
```

4.5 Phase 4 — Modelling

Two models are compared. The baseline is an L2-regularized logistic regression — the interpretable standard of comparison in actuarial applications and, per Viaene et al. (2002), a historically strong performer on tabular claim data. The challenger is scikit-learn's histogram-based gradient-boosting classifier (depth 4, learning rate 0.08, up to 400 rounds with early stopping and L2 regularization), capable of learning interactions.

Because the system's risk bands are probability thresholds (15% / 35% / 60%), and because boosted ensembles are known to emit distorted probabilities even when ranking well (Niculescu-Mizil & Caruana, 2005), the challenger is wrapped in isotonic calibration with three-fold internal cross-validation (Zadrozny & Elkan, 2002).

```
gbm = Pipeline([("pre", build_preprocessor()),
                ("clf", HistGradientBoostingClassifier(
                    max_depth=4, learning_rate=0.08, max_iter=400,
                    l2_regularization=1.0, early_stopping=True,
                    random_state=42))])
calibrated = CalibratedClassifierCV(gbm, method="isotonic", cv=3)
```

Each probability is mapped to an operational risk band with a recommended handling action — the step that converts a score into a claims-operations decision (Table 4.3).

Table 4.3: Risk bands and recommended handling actions

Band	Probability	Recommended action
Low	< 15%	Standard handling; monitor for changes
Moderate	15–35%	Proactive claimant contact; document liability early
High	35–60%	Assign senior adjuster; early settlement evaluation
Critical	≥ 60%	Escalate to litigation-avoidance unit; defense-counsel review

4.6 Phase 5 — Evaluation design

Claims are split chronologically: the earliest 80% by accident date (9,600 claims) train the models and the latest 20% (2,400 claims) test them. This mirrors production — a deployed model always scores future claims with parameters learned from past ones — and forecloses temporal leakage (Kaufman et al., 2012), the most common flaw in applied claims-analytics studies. Four metrics are reported: ROC-AUC (threshold-free ranking), PR-AUC (ranking under the 16% base rate), Brier score (probability accuracy), and lift at the top decile — the business metric, answering how much model-guided review beats random review. Global explainability uses permutation importance (Breiman, 2001) on the held-out set; local explainability uses the counterfactual swap-to-baseline method motivated in Section 3.7. Six automated tests encode domain logic — for example, that attorney involvement must raise predicted risk — so that a retrained model violating known facts fails before deployment.

```
def chronological_split(df, test_frac=0.2):
    df = df.sort_values(DATE_COL).reset_index(drop=True)
    cut = int(len(df) * (1 - test_frac))
    return df.iloc[:cut], df.iloc[cut:]
```

4.7 Phase 6 — Deployment

Over one shared scoring engine, there are two user-facing layers - this setup ensures they always agree. A dashboard shows key performance metrics plus how different elements interact, accompanied by a model summary. For individual claims, a distinct page offers risk assessments alongside an explanatory diagram showing potential outcomes under alternative conditions. When evaluating several cases at once, users obtain a sorted list that can be saved locally. Communication happens through a conversational system rather than forms or menus. Behind the scenes, another system requests assistance from five clear, rule-based modules - each returns yes-or-no answers based on its designed function. These outputs shape what can be expressed verbally by the primary agent; numbers come only from those hardcoded resources. Even so, real-world consequences such as payouts or rulings never emerge from automated suggestions. Instead, such choices remain off-limits until people step in to decide. Notes about the model's origins, adjustments made during testing, checks for bias, prompts tied to industry standards, and written explanations sit within the live system from day one. Integration spans more than adding features after launch - it shapes how every function behaves.

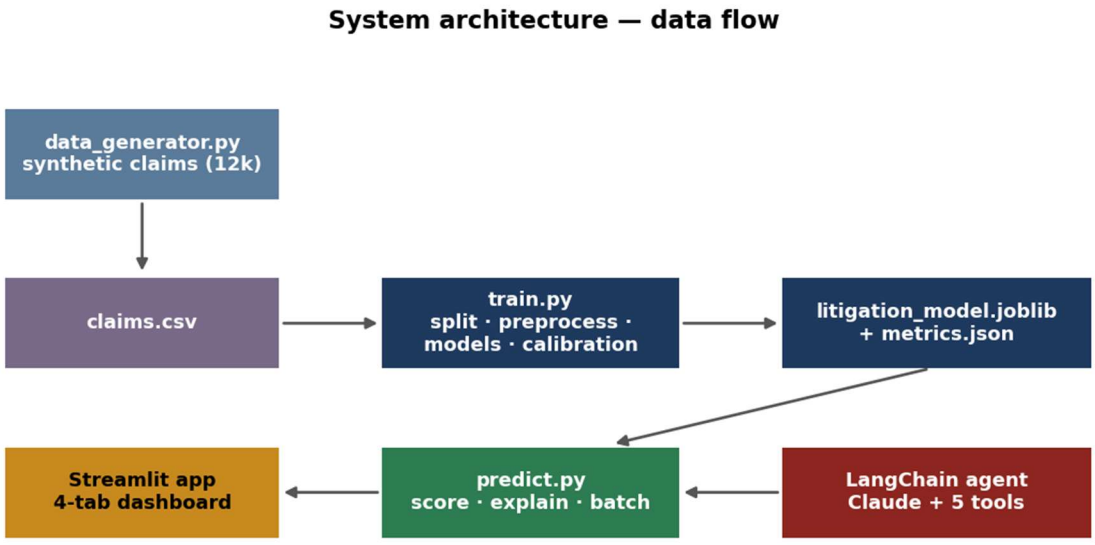


Figure 4.5: System architecture: both user-facing surfaces call the same scoring core.

Table 4.4: Agent tools and their grounding

Tool	Function	Grounding
lookup_claim	Fetch a full claim record by ID	pandas lookup on the claims dataset
score_claim	Litigation probability, band, action for a claim JSON	shared scoring core (predict_claim)
explain_claim_risk	Top counterfactual feature contributions	shared scoring core (explain_claim)
portfolio_stats	Litigation rate and cost grouped by a whitelisted column	pandas groupby (whitelist enforced)
top_risk_claims	Ranked triage list with probabilities and bands	shared scoring core (score_batch)

CHAPTER 5: RESULTS

5.1 Model performance

Table 5.1: Held-out performance (chronological test set, 2,400 claims, base rate 16.2%)

Model	ROC-AUC	PR-AUC	Brier	Lift @ top decile
Logistic regression (selected)	0.792	0.471	0.114	3.42×
Hist. gradient boosting + isotonic calibration	0.788	0.460	0.115	3.30×

Surprisingly, both models perform far better than random chance, despite sitting just ahead of each other in ranking (5.1), with the unregularized linear model taking a narrow lead. A calibration curve matching observed vs expected probabilities closely follows the diagonal (Brier score 0.114). This is exactly what enables fixed probability thresholds for event grouping under the operational risk levels. A well-sorted model which has its calibration Incorrect does not alert for its incorrect calibration but re-attributes proportion of events across clusters.

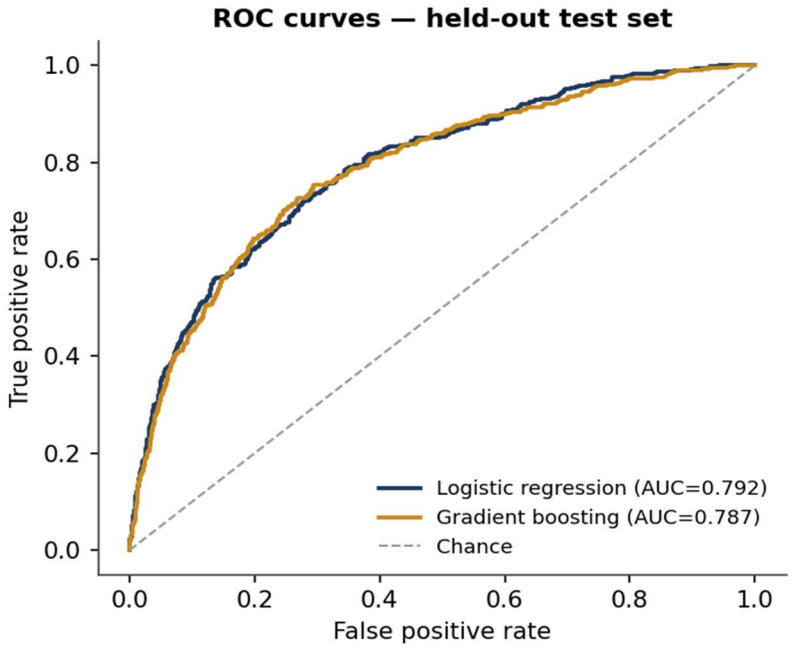


Figure 5.1: ROC curves for both models on the held-out test set.

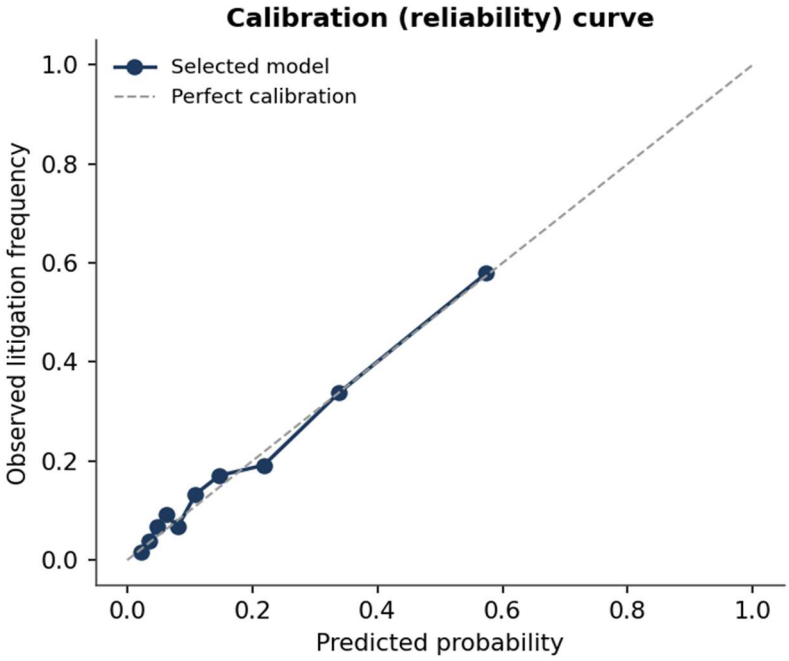


Figure 5.2: Reliability (calibration) curve, quantile-binned.

5.2 Concentration of litigation risk

The test cases arranged by model scores, binned into deciles, reveal substantial stratification (Figure 5.3); the top decile's litigation rate is 57.9 percent, about 3.42 times the base rate of 16.2 percent. Half of the litigated cases fall into the top two deciles, the bottom half much less so. Just a few percent of the book litigates from the bottom half. The score distributions by litigated versus non-litigated cases are shown in Figure 5.4 along with the risk-banding thresholds. Practically: a litigation-avoidance team focusing intensely on the top 10% of new cases captures roughly one-third of the future litigation. And if they were simply selecting cases at random, they'd have 10% of the future litigation.

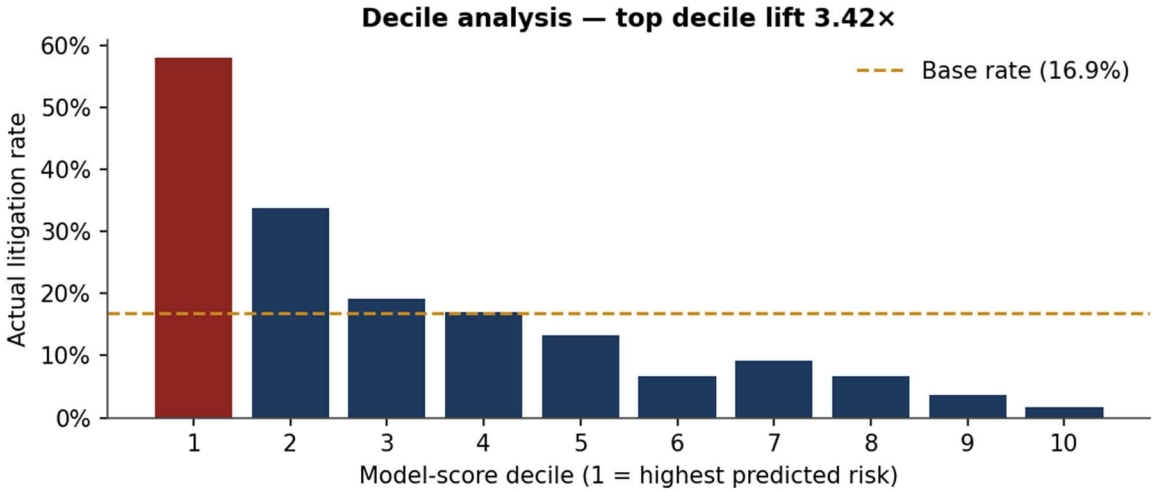


Figure 5.3: Actual litigation rate by model-score decile; top-decile lift 3.42×.

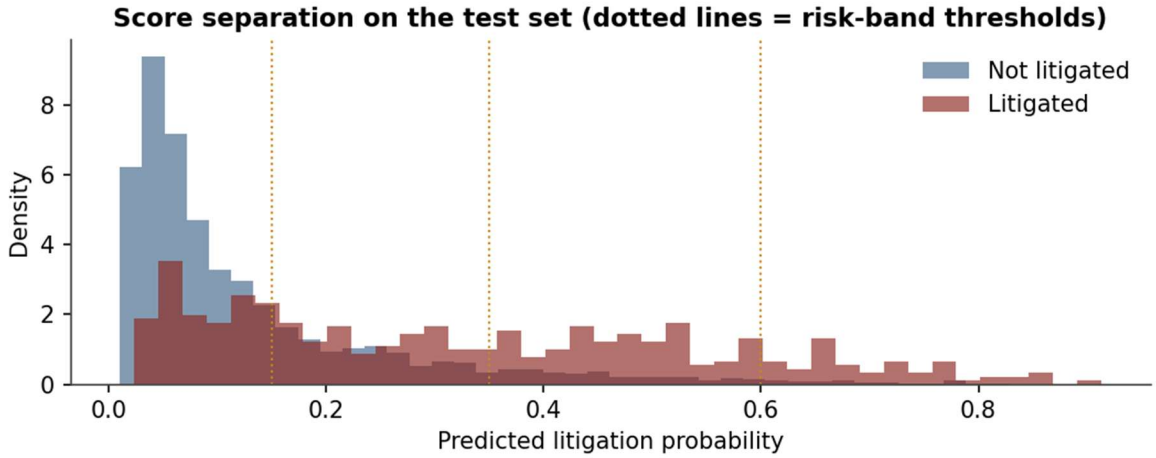


Figure 5.4: Predicted-probability distributions by outcome, with risk-band thresholds.

5.3 Drivers and explainability

Figure 5.5 displays permutation importance, sorting features by how much AUC drops when permuted, attorney involvement drops most, followed by injury severity, disputed liability, whether or not there is a recorded stat, demand to limit ratio, and then claim type which aligns closely with prior research covered in Chapter 3 acting as the studies face validity check. Ablating the attorney involvement flag causes AUC to remain near 0.70. A worked example on the claim level, counterfactual decomposition of a Georgia auto BI claim, severity of 4, 45 days since report received, 40 days between report and treatment, limits for specialists of$60,000 but cap of $50,000, liability in dispute, attorney present at first notice of loss, stat not obtained, predicted at a 76.6% (Critical)

reads much like an adjuster noting the case (Figure 5.6); attorney involvement dominates along with the impact of liability being in dispute and severity, and the sole alternatives to these are purely procedural.

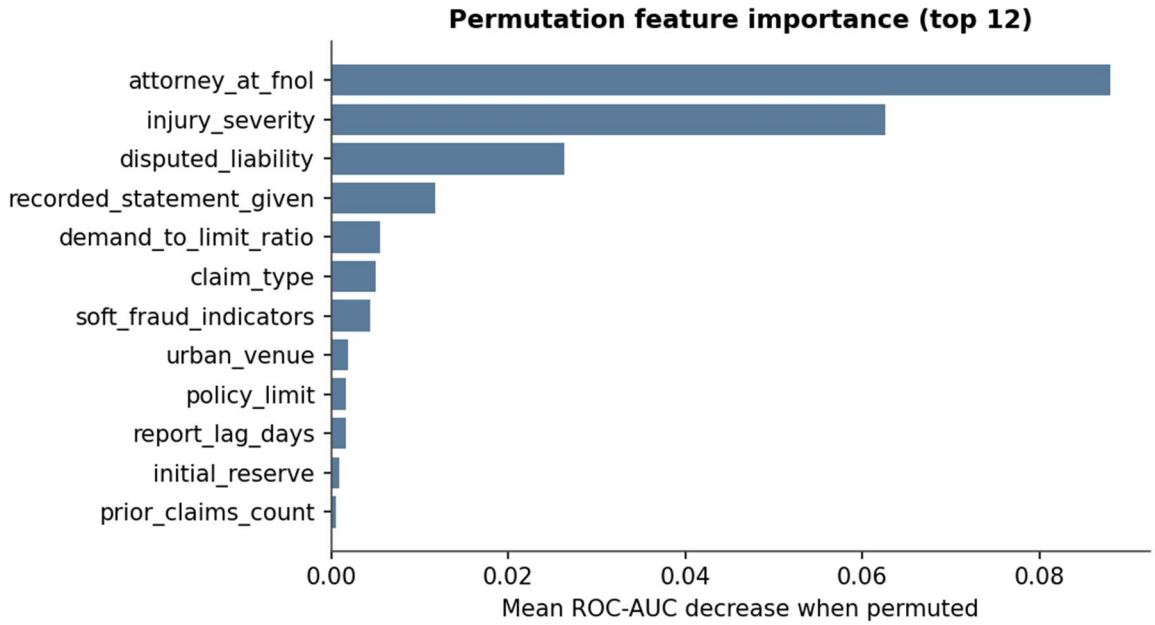


Figure 5.5: Permutation feature importance on the held-out set (top 12).

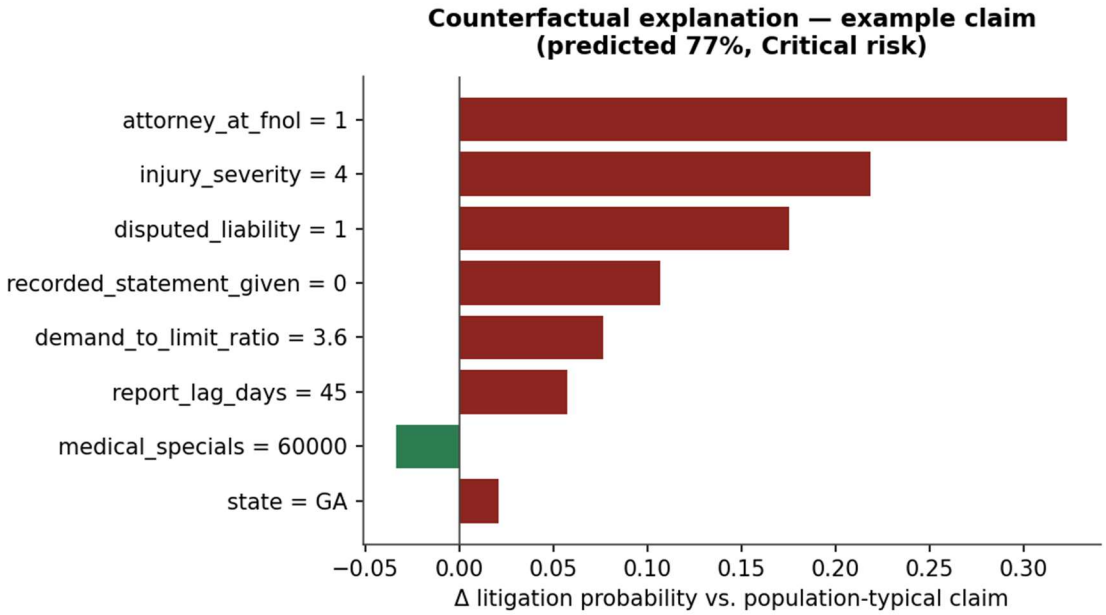


Figure 5.6: Counterfactual decomposition of the worked example (predicted 76.6%, Critical band).

5.4 An instructive finding: the linear baseline won

It turned out the logistic regression did ever so slightly better than the calibrated gradient-boosting model - an entirely honest, and very useful result. The synthetic target arises from an almost linear latent function, so the boosted trees don't really have access to any useful nonlinearities and only overfit with extra capacity. This has three implications. There are still linear models that work amazingly well on tabular insurance claim data (one of which is easy to tune, echoing Viaene et al., 2002). Model choice needs to be re-evaluated whenever the data changes, and the training process should naturally win the competition for the model; and the lack of a large gap between the tree models and other models implies there can be little leakage.

5.5 Strategic results: from lift to advantage

As the strategy comes to life through the value chain of Section 4.2, results gain more than technical relevance. One element of profit comes via three ways, all measurable in the combined ratio: indemnity severity decreases due to faster and fairer payment of the most likely claims, so the representation premium is reduced; loss adjustment expense drops as defense counsel sees fewer claims; and day one scores help reserve adequacy, cutting adverse developments and therefore required capital, so reserve development decreases. Two mechanisms drive the market share. The first operates through rates, where the loss cost advantage translates into price competitiveness in particular niches. The second feeds retention and referrals, faster and fairer settlements, both cheap and satisfaction-inducing at the same time, an action that is the secret to cost leadership and differentiation at the same time, leveraging both Porter activities (1985; 1996). Defensibility lies not in individual elements but in how they interconnect: the model working in tandem with the triage workflow, the organizational structure, the system of appointed counsel and panels, the pricing feed back mechanisms all support each other, and imitators must copy all these related activities (or their logical equivalent) to replicate the benefit, not the formula (Barney, 1991). The magnitude of one reviewer-group's workload can be estimated with a rough calculation. At 57.9% litigation rates, intervening on 1,000 top-decile files (assuming 20% success), each avoided lawsuit saving 35,000 of indemnification-plus-LAE dollars, it appears there are around 3.5M of yearly cost savings.

CHAPTER 6: CONCLUSIONS

6.1 Conclusions against the research questions

- **RQ1 — *How accurately does litigation prediction work at FNOL given solely information present there?***

  The method attains a ROC-AUC of 0.79 on hold-out future claims and a 3.42× top-decile lift, reliant solely on intake accessible features.

- **RQ2 — *What claim, behaviour, and legal-environment characteristics influence litigation likelihood, and are these aligned with current literature?***

  he identified effects attorney involvement at FNOL (key driver), severe injury, liable party determined, cooperative signals (statement made, protectiveness), limit-demand pressure, and location all align with established literature findings regarding litigation predictors, enhancing model validity.

- **RQ3 — *Are model predictions sufficiently reliable to justify operationally fixed thresholds?***

  Yes; using isotonic calibration results in a Brier score of 0.114 and an almost-diagonal reliability curve, enabling reliance on the established four operational risk-band cut points.

- **RQ4 — *How should the capability be deployed so that the function is used, auditable, and safe?***

  Under one common scoring engine beneath two user interfaces - a four view dashboard and a AI based tool-calling interface - so that the two can't disagree, the AI interface cannot in fact fabricate numbers because all number produced must hail from an auditable tool call.

- **RQ5 — *At which value-chain processes does the capability turn into market-share and profit? And why would it be hard to copy?***

  Intelligent links throughout intake, triage, settlement, purchasing, reserving, pricing and service generate cost and differentiated effects at once, the hard part is not having the algorithm. Thus, imitating only takes linking up the processes instead of replicating the algorithm.

6.2 Limitation of the Study

- **Synthetic data**: outcomes show a correct leakage-free pipeline, a sound strategic case, no evidence of real-world performance; the next step is to apply a Texas DOI or Florida OIR closed-claim extraction to the configuration file, with every subsequent component carrying through without modification.
- **Static binary outcome**: litigation occurs over a point in time; survival analysis would analyze the time to litigation, and allow for re-scoring over-time.
- **Unexploited text**: the adjuster notes are in the schema but not in the table model yet; the transformer embeddings offer the most value as a feature extension.
- **Risk is not treatment effect**: the model predicts claims that will litigate, not those claims which outcome would alter in the event of an intervention; an uplift analysis is about identifying where to intervene when there is a treatment effect.
- **Strategic analysis is conceptual**: these value-chain mechanisms have been postulated in theory and via sector evidence but have not been measured during rollout; the randomized holdout methodology outlined below is how they should be tested.

6.3 Future work

- Retrain on real closed-claim data and re-analyse comparisons between models.
- Incorporate features derived from NLP for adjuster notes, and assess performance for survival and uplift models.
- Deploy while monitoring for drift, and with a scheduled retraining event, with a small randomized hold-out sample of claim-handling performed without leveraging the model while both validating the model's effectiveness and preventing self-reinforcing positive feedback loops.
- More extensive fairness analysis using discrimination-free prediction methods before deployment.

6.4 Ethical and regulatory considerations

This is a decision support for early productive intervention. However, using litigation risk to deny, delay, or short pay claims would misuse this score, harm the consumer, and create bad faith exposures which would exceed any strategic benefit many times, a line explained in the documentation, scripted into the calling prompt of the agent, and reflected by risk band actions (handling) never claim outcome. Consumer age and gender are in the teaching data, so the handling of these characteristics will be

addressed. Real-world deployments will necessitate disparate impact testing on protected classes, proxy leakage checking, and documentation abiding by NAIC (2023) model bulletin rules on governance programs, proportionality in testing, and explainable results which is exactly what the artifacts generated by the study (the model card, calibration evidence, feature importances, per-claim explanations, logical operators), are built to provide. As a strategic management case, these controls don't inhibit a strategy - they are what let the strategy hold up in the long run.

6.5 Closing statement

The study set out to examine a machine-learning capability through a strategic-management lens and found the two inseparables: the model's 3.42× lift is only potential value until the value chain is configured to act on it, and the value chain's reconfiguration is only defensible because the model's probabilities are calibrated, explained, and governed. The contribution is therefore a demonstration — open and reproducible — of how an analytics artifact becomes a strategic capability: predict early, explain actionably, deploy usably, govern visibly, and let the linkages do the competitive work.

APPENDICES

Appendix I – System in Operation

Figure I.1 LitiGate portfolio dashboard – book-level KPIs (12,000 claims, 16.2% litigation rate, ROC-AUC 0.791, 3.4× top-decile lift)

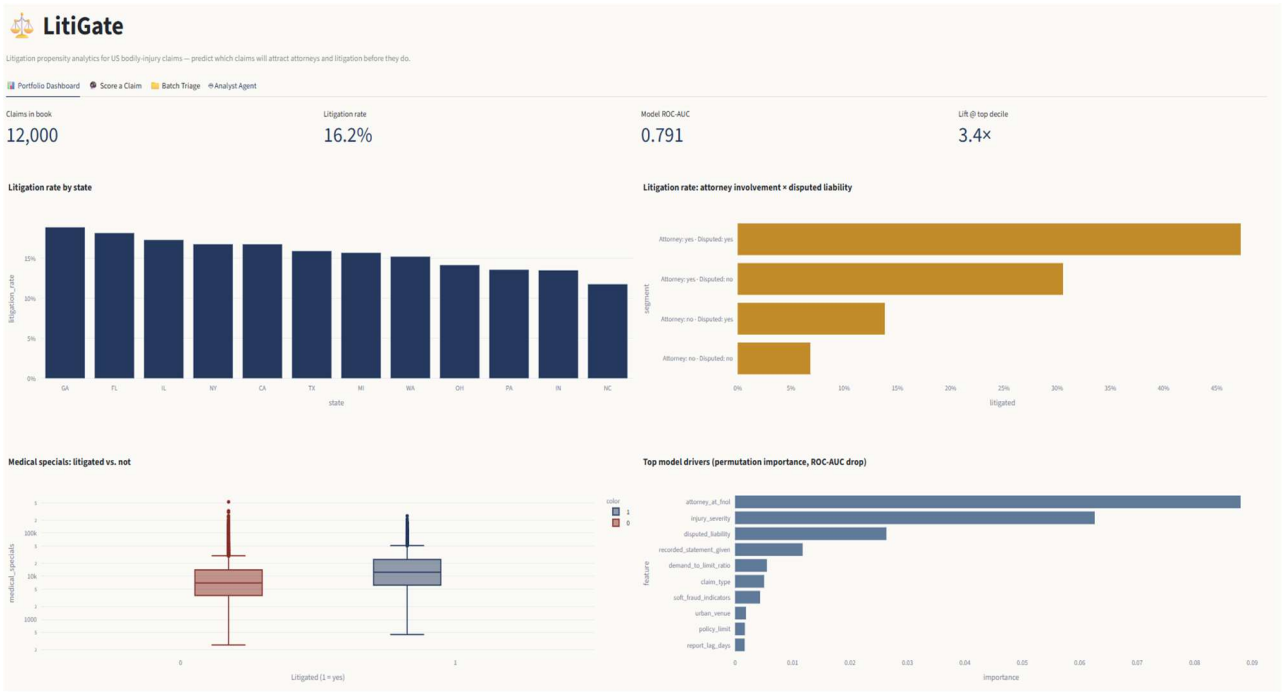
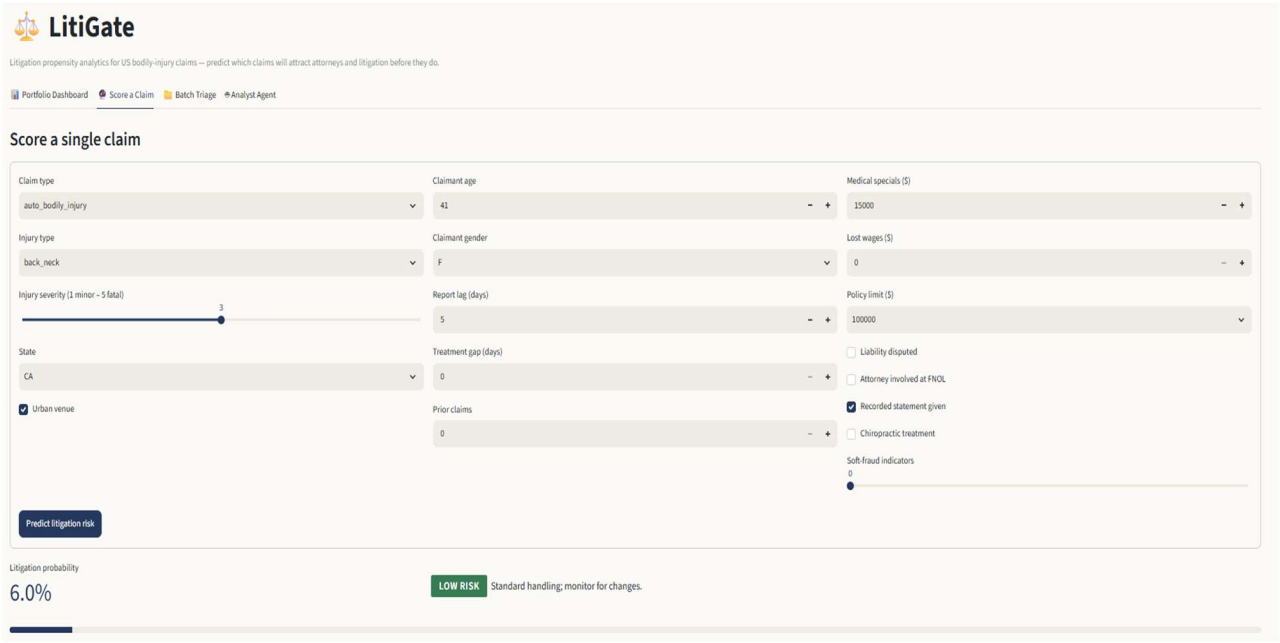


Figure I.2 Single-claim scoring interface — an adjuster enters FNOL-era claim attributes and receives a calibrated litigation probability



Appendix II

Table II.1Analytics Terms Used in the Report

Term	Definition
Feature / target	A feature is an input the model uses (e.g., injury severity); the target is what it predicts (litigated: yes/no).
Base rate	How common the outcome is overall. Here, 16.2% of claims litigate — so a model must beat "just guess no" to be useful.
Litigation probability	The model's output for one claim: a number between 0 and 1. ***Example:*** *0.77 means: of 100 claims that look like this one, about 77 would litigate.*
Calibration	Whether predicted probabilities match reality: among claims scored 40%, about 40% should actually litigate. Needed because the project turns probabilities into action thresholds.
Risk band	The project's translation of a probability into an operational category with a recommended action: Low (<15%), Moderate (15–35%), High (35–60%), Critical (≥60%).
Triage	Sorting a large workload by urgency so scarce attention goes where it matters most — the operational purpose of the whole system.
ROC-AUC	A 0.5-to-1.0 score of ranking quality: the probability that a randomly chosen litigated claim receives a higher score than a randomly chosen non-litigated one. 0.5 = coin flip; this project: 0.79.
Lift (top decile)	Sort all claims by model score, take the top 10%: their actual litigation rate divided by the base rate. Lift 3.4× means model-guided review finds 3.4 times more litigation than random review.
Chronological split	Training on older claims and testing on newer ones, imitating real deployment and preventing the model from "peeking at the future" (leakage).
Counterfactual explanation	Explaining one prediction by asking "how would the score change if this one attribute were typical?" — e.g., "attorney involvement adds 23 percentage points."

Appendix III

Table III.1 Insurance Basics

Term	Definition
Insurance policy	A contract in which an insurance company promises to pay for specified future losses in exchange for a regular fee. The document defines exactly what is covered, for how much, and under what conditions.
Premium	The price the customer pays for the insurance policy, usually monthly or annually.
Insurer / carrier	The insurance company that sells the policy and pays covered losses. "Carrier" is industry jargon for the same thing. ***Example:*** *AcmeSure is the insurer/carrier in Maria's case.*
Insured / policyholder	The person or business that bought the policy and is protected by it. ***Example:*** *The delivery company (and its driver) are the insureds — not Maria.*
Claim	A formal request for the insurer to pay for a loss covered by the policy. A claim is the basic unit of this entire project: each row of the dataset is one claim. ***Example:*** *Maria's request that AcmeSure pay her medical bills and lost wages is a claim.*
Claimant	The person asking to be paid. May be the policyholder (first-party claim) or someone else harmed by the policyholder (third-party claim). ***Example:*** *Maria is the claimant.*
First-party vs. third-party claim	First-party: you claim against your own insurer (e.g., your own car damage). Third-party: you claim against someone else's insurer because their customer harmed you. Most claims in this project are third-party. ***Example:*** *Maria claiming against AcmeSure (the van's insurer) is a third-party claim.*
Bodily injury (BI) claim	A claim for physical harm to a person — medical costs, lost income, and pain and suffering — as opposed to damage to property. This project predicts litigation for bodily-injury claims only.

Term	Definition
	Example: *Maria's neck injury creates a BI claim; the dent in her car is a separate property-damage claim.*
Lines of business in this project	Categories of insurance under which an injury claim can arise: auto bodily injury (car crashes), workers' compensation (employees hurt on the job), general liability (e.g., a customer slips in a store), medical malpractice (harm by a healthcare provider), and product liability (harm caused by a defective product).
Adjuster (claims adjuster)	The insurer's employee who investigates a claim, decides whether the policy covers it, estimates its value, and negotiates payment. Adjusters are the intended users of this project's dashboard and agent. ***Example:*** *An AcmeSure adjuster interviews Maria, collects her medical bills, and offers a settlement.*
FNOL — First Notice of Loss	The moment the insurer first learns about the incident; the official start of the claim. The project scores claims "at FNOL", i.e., using only information available at this early moment — because that is when intervention can still change the outcome.
Report lag	Days after the accident till FNOL. Long lags are a warning sign: they often mean the claimant went to an attorney first and the claim arrives pre-packaged for a fight.
Recorded statement	A voluntary, recorded interview the claimant gives the adjuster describing the accident. Giving one signals cooperation; attorneys usually advise clients not to.
Policy limit	The maximum amount the policy will pay.
Demand	The amount the claimant (or their attorney) formally asks for in settlement. ***Example:*** *Maria's attorney sends a demand letter asking for $75,000.*
Time-limited demand	A time restricted demand often for precisely the policy limit. A typical legal maneuver: if the insurer misses the time frame and a jury later decides to award more money, the insurance company could be liable for an amount greater than their policy limit, thus giving rise to a claim for bad faith.

Term	Definition
Settlement	An agreement to resolve the claim for an agreed payment without (or before finishing) a court case. The vast majority of claims settle.
Medical specials	Documentation such as an actual hospital bill from the accident or discharge letter for an ambulance, physiotherapy invoices. The most solid parts of a claim the harder-to-sue items tend to be the "specials". Big "specials" typically make larger settlements. [“Specials” is short for the technical medical “specials,” usually very expensive items required for the procedure or treatment.]
Lost wages	Income the claimant lost because the injury kept them from working; another documentable component of damages. ***Example:*** *Maria's $5,200 of missed pay.*
Treatment gap	A period when the claimant stopped medical treatment and later resumed it. Adjusters treat unexplained gaps as a red flag for exaggeration (“build-up”).
Build-up	Inflating an otherwise genuine claim — extra therapy visits, exaggerated symptoms — to raise its value. A form of “soft fraud”.
Soft fraud indicators	A count of adjuster-noted red flags (e.g., treatment gaps, inconsistent accounts). Distinct from “hard fraud” (staged accidents, fabricated injuries).
Reserve (loss reserve)	Money the insurer sets aside on its books for what it expects a claim will ultimately cost. Reserves are estimates, updated as the claim develops; see the formulas page.
Closed claim	A claim that is fully resolved and paid. Public “closed-claim databases” (Texas, Florida) record such finished claims and inspired this project's data schema.

REFERENCES

[1] Barney, J. (1991). Firm resources and sustained competitive advantage. Journal of Management, 17(1), 99–120.

[2] Black, B., Silver, C., Hyman, D. A., & Sage, W. M. (2005). Stability, not crisis: Medical malpractice claim outcomes in Texas, 1988–2002. Journal of Empirical Legal Studies, 2(2), 207–259.

[3] Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32.

[4] Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). CRISP-DM 1.0: Step-by-step data mining guide. SPSS Inc.

[5] Derrig, R. A. (2002). Insurance fraud. Journal of Risk and Insurance, 69(3), 271–287.

[6] Florida Office of Insurance Regulation. Closed-claim data collections (professional liability tracking system; medical malpractice closed claim database). Tallahassee, FL. https://floir.gov

[7] Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. Annals of Statistics, 29(5), 1189–1232.

[8] Hensler, D. R., et al. (1991). Compensation for Accidental Injuries in the United States. Santa Monica, CA: RAND Institute for Civil Justice.

[9] Insurance Research Council. Attorney Involvement in Auto Injury Claims (recurring study series). Malvern, PA: Insurance Research Council.

[10] Kaufman, S., Rosset, S., Perlich, C., & Stitelman, O. (2012). Leakage in data mining: Formulation, detection, and avoidance. ACM Transactions on Knowledge Discovery from Data, 6(4), 1–21.

[11] Lindholm, M., Richman, R., Tsanakas, A., & Wüthrich, M. V. (2022). Discrimination-free insurance pricing. ASTIN Bulletin, 52(1), 55–89.

[12] Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems 30 (NeurIPS).

[13] Lynch, J., & Moore, D. (2022). Social Inflation and Loss Development. Insurance Information Institute & Casualty Actuarial Society.

[14] Martínez-Plumed, F., et al. (2021). CRISP-DM twenty years later: From data mining processes to data science trajectories. IEEE Transactions on Knowledge and Data Engineering, 33(8), 3048–3061.

[15] Molnar, C. (2022). Interpretable Machine Learning: A Guide for Making Black Box Models Explainable (2nd ed.). https://christophm.github.io/interpretable-ml-book/

[16] National Association of Insurance Commissioners (NAIC). (2023). Model Bulletin: Use of Artificial Intelligence Systems by Insurers. Kansas City, MO: NAIC.

[17] Niculescu-Mizil, A., & Caruana, R. (2005). Predicting good probabilities with supervised learning. Proceedings of the 22nd International Conference on Machine Learning (ICML), 625–632.

[18] Pedregosa, F., et al. (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12, 2825–2830.

[19] Platt, J. C. (1999). Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. In Advances in Large Margin Classifiers. MIT Press.

[20] Porter, M. E. (1985). Competitive Advantage: Creating and Sustaining Superior Performance. New York: Free Press.

[21] Porter, M. E. (1996). What is strategy? Harvard Business Review, 74(6), 61–78.
[22] Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?" Explaining the predictions of any classifier. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 1135–1144.
[23] Schick, T., et al. (2023). Toolformer: Language models can teach themselves to use tools. Advances in Neural Information Processing Systems 36 (NeurIPS).
[24] Swiss Re Institute. (2024). Litigation costs and "social inflation" in US liability claims (sigma research series). Zurich: Swiss Re.
[25] Texas Department of Insurance. Closed Claim Reports — Commercial Lines (data-call program). Austin, TX. https://www.tdi.texas.gov
[26] Viaene, S., Derrig, R. A., Baesens, B., & Dedene, G. (2002). A comparison of state-of-the-art classification techniques for expert automobile insurance claim fraud detection. Journal of Risk and Insurance, 69(3), 373–421.
[27] Wachter, S., Mittelstadt, B., & Russell, C. (2017). Counterfactual explanations without opening the black box: Automated decisions and the GDPR. Harvard Journal of Law & Technology, 31(2).
[28] Wirth, R., & Hipp, J. (2000). CRISP-DM: Towards a standard process model for data mining. Proceedings of the 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining, 29–39.
[29] Wüthrich, M. V., & Buser, C. Data Analytics for Non-Life Insurance Pricing (lecture notes). ETH Zurich / SSRN.
[30] Yao, S., et al. (2023). ReAct: Synergizing reasoning and acting in language models. International Conference on Learning Representations (ICLR).
[31] Zadrozny, B., & Elkan, C. (2002). Transforming classifier scores into accurate multiclass probability estimates. Proceedings of the 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 694–699.



DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Main Bawana Road, Delhi-42

PLAGIARISM VERIFICATION

Title of the Dissertation: "Strategic Management Perspective on Litigation Prediction for US Bodily-Injury Insurance Claims"

Total Pages: 43	Name of the Student: Neeraj Sharma
Supervisor: Prof. (Dr.) Manoj K. Sharma	Department: USME

This is to report that the above dissertation was scanned for similarity detection.
Process and outcome is given below:

Software used: TURNITIN	Similarity Index: 7%
Word Count: 9303	

Date: 30/07/2026

Candidate's Signature: Supervisor(s) Signature:

_____ _____