Progress of the doctoral project entitled:

Artificial Intelligence techniques for the decision support in credit management

Presented on January 13, 2026

PhD Student: **Omar Serghini**

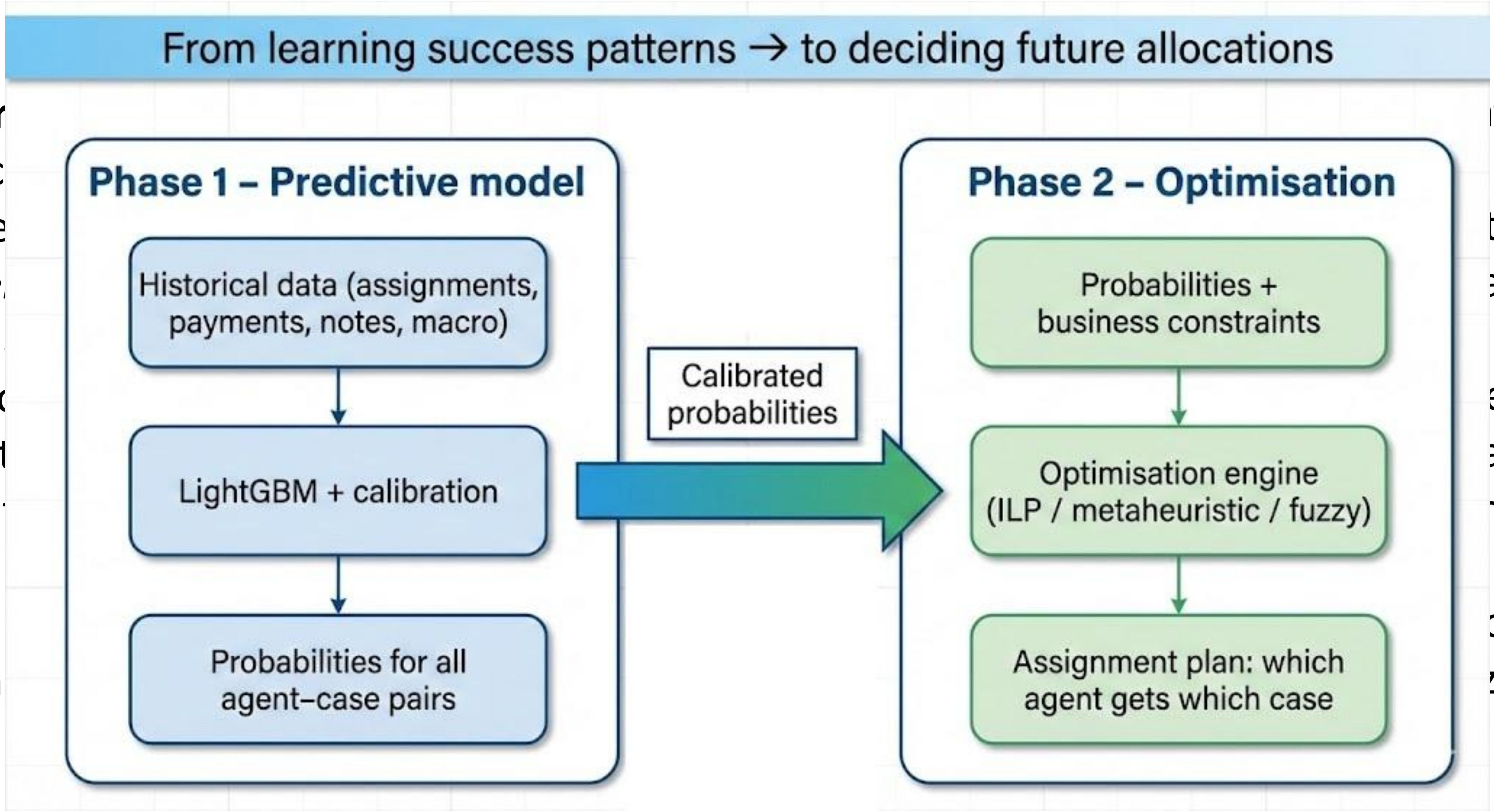
Supervisors: **- Marco Lucio Scarpa**
- Salvatore Serrano

Problem & motivation

- **Context**: European credit agency managing large volumes of post-default utility bills; each file is assigned to an internal agent with a limited recovery window.
- **Current situation:** Case-agent assignment is mostly rule-based and manual (areas, phases, workload), not driven by explicit success probabilities.
- **Why it's difficult**: Utilities debts are low-ticket, high-volume, with limited debtor information, strict timelines, and standardized procedures → hard to see where agent choice really matters.
- **Opportunity**: The agency already logs rich operational data (assignments, payments, contacts, notes, phases). This can be turned into probabilities of success for each agent–case pair.
- **Goal of my work:** Develop and validate a calibrated model that scores all feasible agent-case pairs, as a basis for future optimisation of case distribution under real constraints.

Project overview - two phases

For each chosen real assignment (case–agent pair):



Phase 1 - Pr

- Build a c
estimate
$P(recove$
feasible
- Use histo
payment
indices)
succeed
- Validate
uplift an

tion (next

ties as inputs
at distributes

ed recovery
agent
rness,

on families:
zy / multi-

Data & panel construction

- **Source & scope:**
  - Proprietary FireSpa Italia data on defaulted utilities contracts.
  - Relational tables: assignments, contracts, debtors, payments, notes, agents, mandanti, plus macro-economic series.
- **Unit of analysis:**
  - Each row = one (agent, case, assignment sequence) with a defined start/end date.
  - TARGET = 1 if any payment occurs during that window, 0 otherwise.
  - Reassignments ⇒ same case may appear multiple times with different agents.
- **Time window & split:**
  - Period: Nov 2022 - Nov 2024 (utilities only).
  - Train/validation: up to 2 July 2024; Test: from 3 July 2024 onward.
  - Extra rule: per-case sequence monotony to avoid leakage across the cutoff.
- **Final panel:**
  - After joins, cleaning, and case-history reconstruction: ≈ 746k train/val assignments, ≈ 166k test assignments.

Feature families (high level)

- To keep the model interpretable, all inputs are grouped into semantic families.
- Most signal comes from case/process and behavioral features, not from demographics.
- Sensitive attributes (age, gender/person type, region, agent location) are explicitly flagged and monitored.

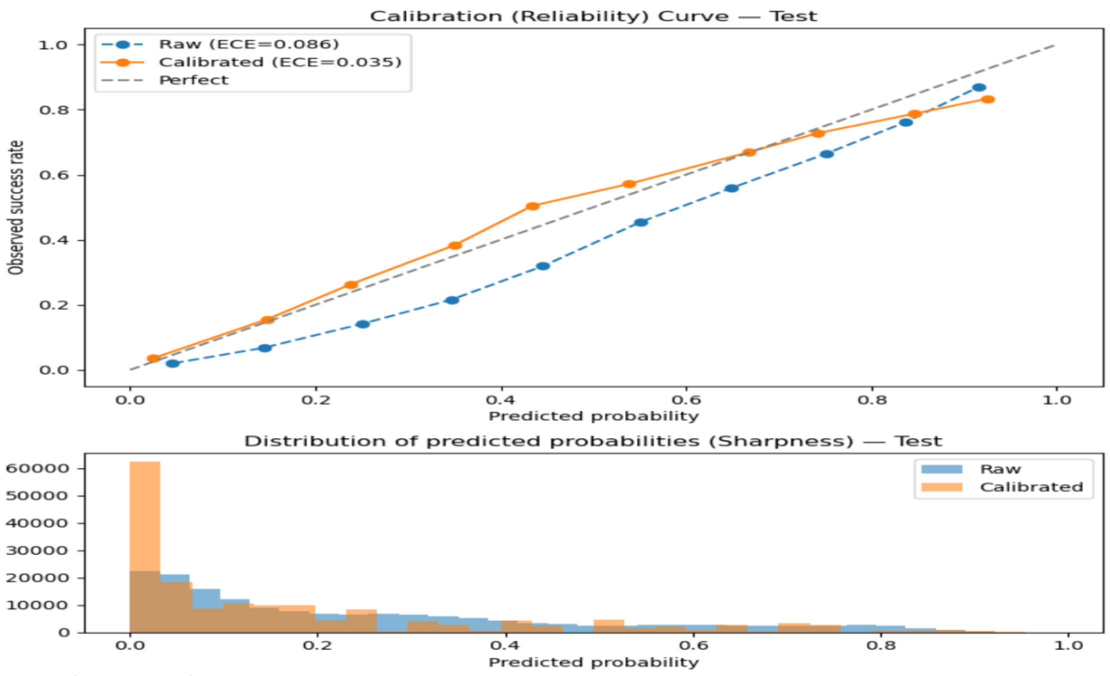
Code	Family	What it captures (very short)
CRI	Case info	Contract, product, amounts, process status
ASG	Assignment	Phase, sequence, dates, mandate context
PRH	Prior hist	What happened on the case before this assignment
ABM	Behav. (agent)	Cumulative moves: calls, promises, legal, etc.
APM	Perf. (agent)	Success rates, recovered amounts, reassignments
DEB	Debtor	Age band, person type (B/F/M), region, segment
ARI	Agent	Branch, province, role, age band
ECO	Macro	GDP, unemployment, income, inflation by quarter

Predictive model setup

- Unit of prediction: one row = (agent, case, assignment sequence) with a defined start/end window.
- Target: 1 if any recovery occurs during that assignment window, 0 otherwise.
- Inputs: ~170 features (case + assignment context, prior case history, agent behaviour & performance, macro indicators).
- Model: LightGBM (gradient-boosted trees), tuned with Optuna on a chrono split (train / validation).
- Training strategy:
    - Train/val before 2 July 2024, test from 3 July 2024 onward (true “future” test).
    - Sample weights up-weight high-amount recoveries so the model cares more about economically relevant wins.

Model performance & calibration

- **Discrimination:**
  - Test AUC ≈ 0.865 on future data (after 3 July 2024).
  - Stable across major utilities and assignment phases
- **Raw probabilities:**
  - Good ranking, but under- / over-confident in some ranges



- **Calibration step:**
  - Compared Platt scaling vs isotonic regression on the validation window.
  - Chosen: isotonic, as it reduced Brier score and calibration error without changing AUC.
- **Outcome:**
  - Post-calibration probabilities are trustworthy: a score of 0.30 now means ≈30% success on similar files.
  - This is essential because phase 2 will use these numbers directly in the assignment optimizer.

Simulation of assignment scenarios

What we simulate

- For a subset of test cases, we create synthetic agent–case pairs and score them with the calibrated model.
- This lets us compare:
  - ✓ Uplift = best (top-5) predicted agent vs median agent for the same case.
  - ✓ Rank hit-rate = where the historical agent appears in the model's ranking.

Key findings (high level)

- Most cases show small but positive uplift (≈1–5 percentage points).
- A non-trivial share exhibits larger gains, meaning agent choice can matter.
- Historical assignments are not random: the real agent often appears among the better candidates, but not systematically at the top.

Metric	Value
Cases with spread ≥ 3 pp	46.9%
Cases with spread ≥ 5 pp	38.5%
Cases with uplift ≥ 5 pp	30.8%
Median spread (best – worst agent)	1.82 pp
Median uplift (top-5 – median)	1.46 pp
Median rank of actual agent	21

Predictive model setup

- Sensitive features in scope:
  - Age, gender/person type, region/province (debtor & agent) are explicitly flagged.
  - Together they explain only ≈3-4% of model gain; most signal comes from case, history, and behavioral features.

- Group-wise performance:
  - AUC and error checked by region, agent age band, debtor type (F, M, Business).
  - Metrics are broadly similar across groups; residual gaps mostly reflect portfolio mix and sample size.

- Global & local explanations (SHAP):
  - SHAP shows top drivers are case/contract, process phase, entrusted amount, prior history, agent behavior.
  - Sensitive features appear low in the ranking with small SHAP impact.
  - For any assignment we can provide a short list of key factors behind the predicted probability, supporting audit and monitoring.

Phase 2: optimization layer

- **From scores to decisions**
  - The model now gives a probability of success for each (agent, case).
  - Phase 2 goal: use these probabilities + business rules to decide *who* should get *which* cases.

- **Conceptual formulation:** Assignment problem: choose $x_{ij} \in \{0,1\}$ (assign / not assign) to maximize $\sum_{ij} p_{ij} \cdot w_j \cdot x_{ij}$ under constraints: capacity, portfolio rules, fairness, territories, SLAs.

- **Candidate approaches under study**
  - Exact ILP – transparent, optimal under clear constraints.
  - Metaheuristics – more flexible for messy/soft rules, scalable to large batches.
  - Fuzzy / multi-objective – explicitly trade off recovery vs workload vs fairness.

Thank you for your attention