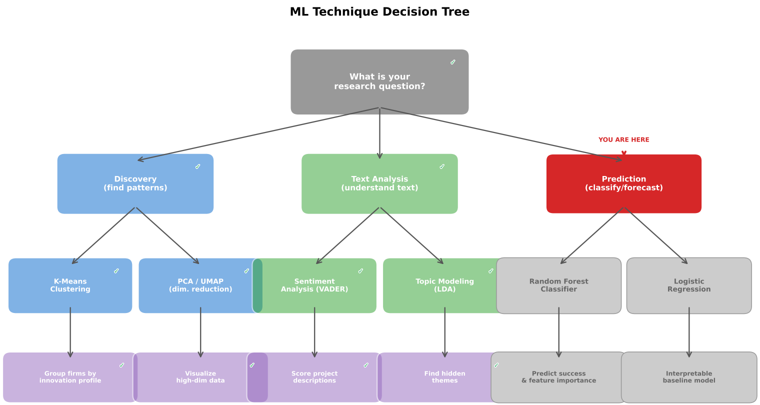


Your research question determines your method. We are entering the Prediction branch.

Session 1 Recap: What We Found

Our 4 innovator clusters:

1. Cautious Incrementors
   Low R&D, low digital, incremental
2. Digital Transformers
   High digital maturity, collaborative
3. Radical Disruptors
   High R&D, VC-funded, radical novelty
4. Collaborative Researchers
   High collaboration, grant-funded

Our NLP findings:

- Sentiment varies by industry
- 5 innovation themes emerged from LDA
- Digital Transformers use different language than Cautious Incrementors

Today's question

Session 1: "What types exist?"
Today: "Can we predict who succeeds?"

Reviewing our Discovery and Text branch findings. Everything from Session 1 is still here. But today, the question changes.

From Description to Prediction

Discovery: What patterns exist?

$\downarrow$

Prediction: What will happen?

$\downarrow$

Generation: Can AI create insights?

This progression mirrors the evolution of data science: describe $\rightarrow$ predict $\rightarrow$ generate.

Supervised Learning: Learning from Labeled Data

The analogy:

A junior researcher reads 300 labeled project reports ("success" / "failure"). After enough examples, they can predict the outcome of a new report.

That's supervised learning:

- Input: features + known outcomes
- Algorithm learns the pattern
- Predicts outcomes for new data

Classification vs. Regression:

- **Classification:** success/failure (categories)
- **Regression:** success score 0–100 (continuous)

For innovation research:

- Predict innovation outcomes
- Identify success factors
- Feature importance = research findings
- Compare what matters vs. what doesn't

Supervised learning needs labeled data. The quality of your labels determines the quality of your predictions.

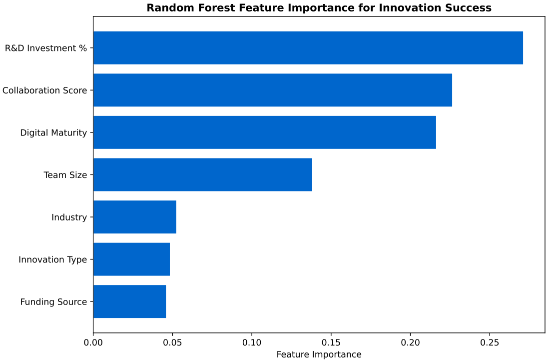
Feature Importance: The Real Research Finding

For researchers, prediction accuracy is secondary.

The real insight is: *which features matter?*

Feature importance tells you:

- R&D investment drives success? How much?
- Does collaboration matter more than team size?
- Is digital maturity a prerequisite?



Feature importance $\neq$ causality. A feature can be important for prediction without causing the outcome.

Your Turn: Where Does This Fit?

"Prediction sounds powerful.
When is it dangerous? When should we NOT predict?"

Take 60 seconds. Discuss with your neighbor.

Map Check: Where Are We?

Branches explored:

- ✓ Session 1: Discovery + Text branches complete
- ✓ Feature importance previewed

Next on the map:

- Classification: building the predictive models
- Random Forest, Logistic Regression, evaluation

We know what matters. Now we build the models and prove they work on data they have never seen.