Innovation Diamond

10-Slide Technical Summary

Machine Learning for Smarter Innovation

BSc Course – Student Reference

The Innovation Diamond: Framework Overview

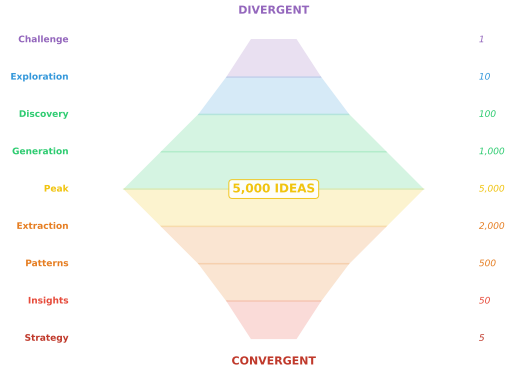
Core Concept:

- **Divergent**: Expand possibilities
- **Peak**: Maximum exploration
- **Convergent**: Focus on best

The Journey:

$$1 \longrightarrow 5{,}000 \longrightarrow 5$$

Challenge $\rightarrow$ *Ideas* $\rightarrow$ *Strategies*



ML enables both creative expansion (generative) and strategic focus (discriminative)

Divergent Phase: Expansion Techniques

ML Foundations – Problem Framing

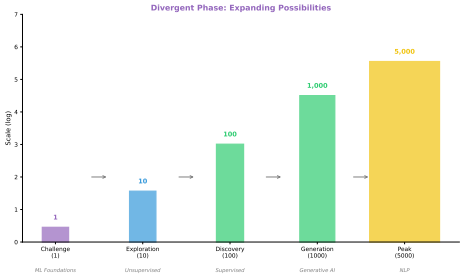
$$L(\theta) = \frac{1}{n} \sum_{i=1}^{n} \ell(y_i, f(x_i; \theta))$$

Unsupervised Learning – Pattern Discovery

$$\underset{C}{\operatorname{argmin}} \sum_{k} \sum_{x \in C_k} \|x - \mu_k\|^2$$

Supervised Learning – Feature Engineering

$$\hat{y} = \sum_{j=1}^{p} \beta_j x_j + \epsilon$$



Stage	Count
Challenge	1
Exploration	10
Discovery	100

Divergent techniques expand from one challenge to hundreds of features and dimensions

Generative AI & NLP: Creative Expansion

Generative AI – Idea Generation (1,000)

$$P(x) = \int P(x|z)P(z)dz$$

LLMs generate diverse solution variants

Topic Modeling – Theme Discovery

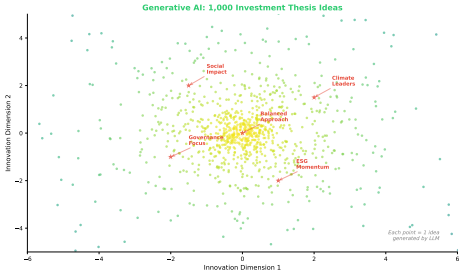
$$P(w|d) = \sum_{t} P(w|t)P(t|d)$$

LDA extracts latent topics from documents

NLP & Sentiment – Text Processing

$$P(w_t|w_{t-k}, \ldots, w_{t-1}) = \text{softmax}(W \cdot h_t)$$

Language models process massive text data



Stage	Count
Generation	1,000
Peak	5,000

Generative techniques expand to thousands of possibilities beyond human capacity

Peak: Maximum Exploration (5,000 Possibilities)

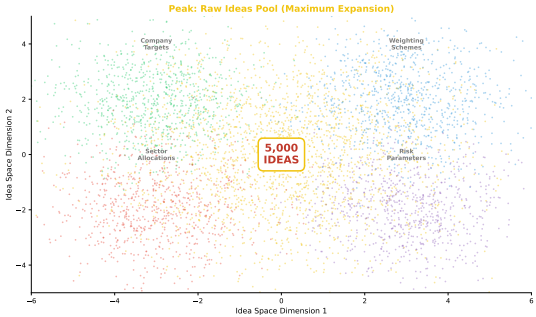
How 5,000 Emerges:

- Companies: 500+ targets
- Weighting schemes: 100+
- Sectors: 50+
- Risk parameters: 20+

Expansion Techniques Used:

- Unsupervised Learning
- Generative AI (LLMs)
- NLP & Sentiment Analysis
- Topic Modeling

Pitfall: Analysis paralysis



The peak represents maximum divergence – now convergent ML techniques must focus

Convergent Phase: Focus Techniques

Clustering – Grouping (5,000 → 2,000)

$$\text{Silhouette}(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))}$$

Group similar ideas, remove duplicates

Classification – Ranking (2,000 → 500)

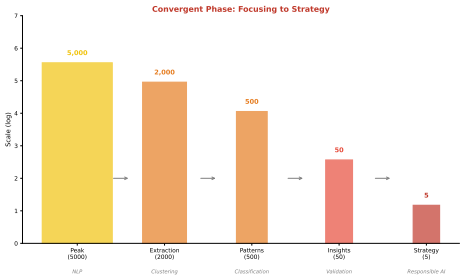
$$\text{Gini}(D) = 1 - \sum_{k=1}^{K} p_k^2$$

Classify by feasibility and impact tiers

Neural Networks – Cross-cutting

$$a^{(l)} = \sigma(W^{(l)} a^{(l-1)} + b^{(l)})$$

Deep patterns across both phases



Stage	Count
Extraction	2,000
Patterns	500

Convergent techniques systematically reduce possibilities to actionable strategies

Validation & Final Selection

Validation – Testing (500 → 50)

$$\text{CV} = \frac{1}{k}\sum_{i=1}^{k} \text{Score}(f_{-i}, D_i)$$

Cross-validation ensures generalization

A/B Testing – Comparison

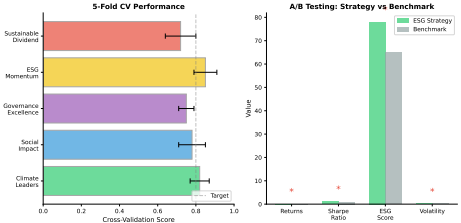
$$t = \frac{\bar{x}_A - \bar{x}_B}{\sqrt{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}}}$$

Statistical significance for decisions

Responsible AI – Selection (50 → 5)

$$\phi_j = \sum_{S} \frac{|S|!(n - |S| - 1)!}{n!}[v(S \cup j) - v(S)]$$

SHAP values for explainability



Stage	Count
Insights	50
Strategy	5

Validation and explainability ensure final strategies are trustworthy and actionable

ESG Case Study: Complete Journey

The Challenge:
"Create a portfolio maximizing returns with genuine ESG impact"

Journey Numbers:

- **1** ESG portfolio challenge
- **10** sustainability dimensions
- **100** engineered features
- **1,000** LLM-generated theses
- **5,000** raw investment criteria
- **2,000** clustered approaches
- **500** classified patterns
- **50** validated insights
- **5** final strategies

Final 5 Strategies:

1. Climate Leaders Fund
2. Social Impact Blend
3. Governance Excellence
4. ESG Momentum Strategy
5. Sustainable Dividend Growth

All 14 ML techniques applied systematically through the Innovation Diamond

ML Toolkit Reference: All 14 Techniques

Topic	Phase	Signature Equation	Purpose
ML Foundations	Divergent	$L(\theta) = \frac{1}{n} \sum \ell(y_i, f(x_i; \theta))$	Problem framing
Supervised	Divergent	$\hat{y} = X\beta + \epsilon$	Feature engineering
Unsupervised	Divergent	$\text{argmin} \sum \|x - \mu_k\|^2$	Pattern discovery
Neural Networks	Both	$a^{(l)} = \sigma(W^{(l)} a^{(l-1)} + b^{(l)})$	Complex patterns
Generative AI	Divergent	$P(x) = \int P(x\|z)P(z)dz$	Idea generation
NLP & Sentiment	Divergent	$P(w_t \| w_{<t}) = \text{softmax}(Wh_t)$	Text processing
Topic Modeling	Divergent	$P(w\|d) = \sum_t P(w\|t)P(t\|d)$	Theme extraction
Clustering	Convergent	$\frac{b(i)-a(i)}{\max(a(i),b(i))}$	Grouping
Classification	Convergent	$\text{Gini} = 1 - \sum p_k^2$	Categorization
Validation	Convergent	$\text{CV} = \frac{1}{k} \sum \text{Score}$	Testing
A/B Testing	Convergent	$t = \frac{\bar{x}_A - \bar{x}_B}{s}$	Comparison
Responsible AI	Convergent	SHAP ϕ_j	Explainability
Structured Output	Convergent	JSON schema	Reliability
Finance	Convergent	VaR_α	Risk modeling

14 ML tools for the complete innovation journey from challenge to strategy

Key Takeaways & Decision Framework

5 Key Insights:

1. ML **amplifies** human innovation
2. **Both phases** are essential
3. **Match technique** to phase
4. **Watch for pitfalls** at each stage
5. **Trust but verify** with validation

Common Pitfalls:

- Divergent: Analysis paralysis, noise over signal
- Convergent: Premature closure, overfitting

Decision Framework:

- Need to explore? → Unsupervised
- Need to predict? → Supervised
- Need to generate? → Generative AI
- Need to group? → Clustering
- Need to rank? → Classification
- Need to test? → A/B Testing

$$\mathbf{1 \rightarrow 5{,}000 \rightarrow 5}$$

The Innovation Diamond: ML-powered systematic innovation from challenge to strategy