

Business Data Analysis

MGM509

Lecture 6

What Makes a Good Model? Decision-Analytic Thinking in Practice

**Lecture
6**

What we will cover

1. What makes a model “good” in real-world deployment
2. Performance measures beyond simple accuracy
3. Cost-sensitive evaluation and decision thresholds
4. Practical examples of decision-focused modeling
5. Summary tools to assess model usefulness

Lecture
6

Data Mining Process Overview (CRISP-DM)

- Cross Industry Standard Process for Data Mining
- Not strictly sequential; often involves **loops and iteration**
- **Evaluating Model Performance is done during Evaluation Phase**

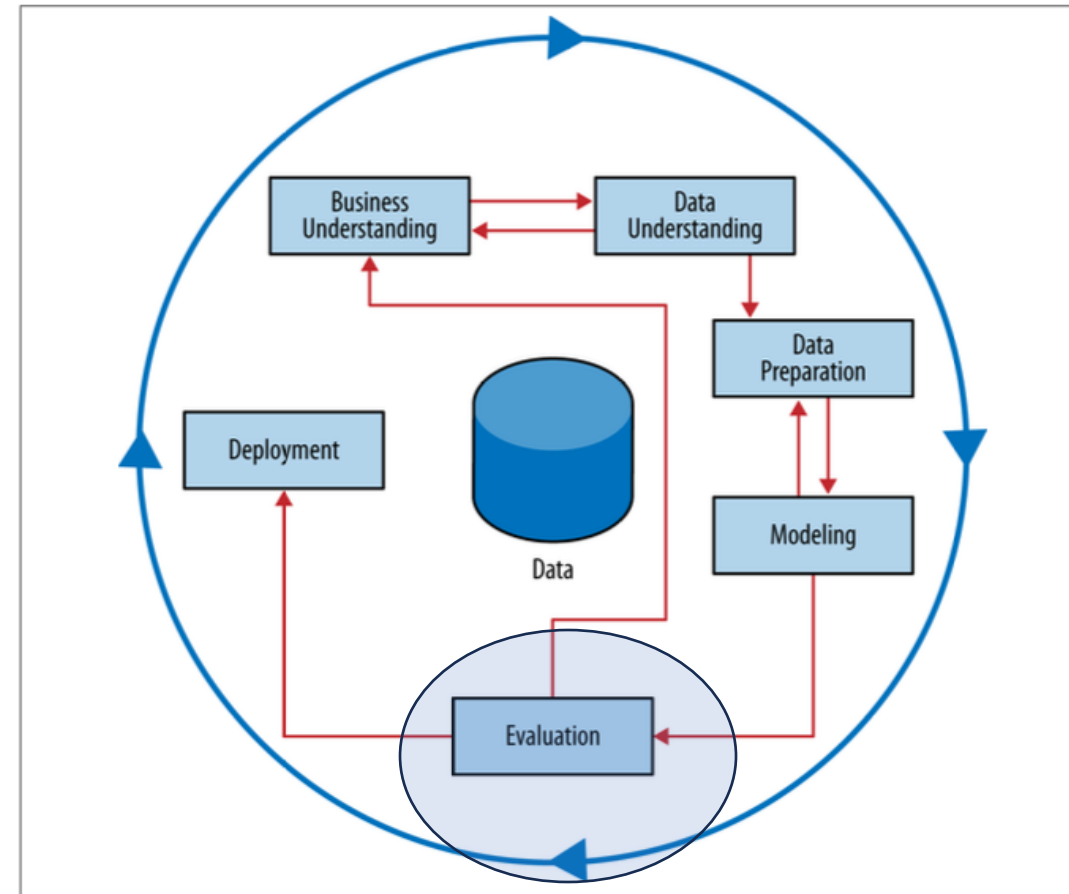


Figure 2-2. The CRISP data mining process.

Lecture
06

Decision Analytic Thinking, Performance Measures, and Cost Sensitive Evaluation



Lecture
6

What Is Decision-Analytic Thinking?

- It's about viewing modeling through a business decision lens, not just a statistical one.
- A technically accurate model may still be useless if it doesn't help the business act effectively.

Core question: “Does this model help me make better decisions?”

It's not about perfect predictions — it's about better outcomes.

Lecture
6

Modeling for Business Utility

Business objectives may include:

- Reducing false alarms (cost-saving)
- Catching rare but critical events (fraud, churn)
- Maximizing revenue or minimizing loss

Decision Analytic Thinking asks:

- What happens if we act on this prediction?
- What are the costs and benefits of correct vs. incorrect decisions?

Lecture 6

Predictive Modeling in Business Context

Model Focus	Traditional Data Science	Decision Analytics
Objective	Maximize accuracy	Maximize business value
Evaluation metric	RMSE, Accuracy, AUC	Expected cost/benefit
Output	Prediction	Action or decision guidance

- For example, a model with slightly lower accuracy but higher net profit (after marketing cost) is a better model.

**Lecture
6****Model Performance in Context**

- Performance must be contextualized by the value of correct and incorrect outcomes. Always ask: “What decisions will this model influence?”
 - Predicting fraud → avoids loss
 - Predicting churn → improves retention strategy
 - Predicting sales → informs inventory and staffing

**Decision
Flow**

Model Prediction

Business Action

Financial Outcome

Lecture
6

Why Accuracy Isn't Always Enough?

- Real business value comes from correctly identifying the actionable minority (e.g., true churners).
- $\text{Accuracy} = \frac{\text{Number of Correct decisions made}}{\text{Total number of decisions made}}$
- Skewed or imbalanced data is when one output is more dominant than the other. For ex. Churn Rate is 10% and Non-Churn is 90%
- Hypothetically, if your model does nothing and just assumed all customer will Not Churn → you will have 90% accuracy.

**Lecture
6****Beyond Accuracy – The Confusion Matrix**

- The confusion matrix provides a foundation for all other evaluation metrics:

	Predicted: Positive	Predicted: Negative
Actual: Positive	True Positive (TP)	False Negative (FN)
Actual: Negative	False Positive (FP)	True Negative (TN)

- Positive = will Churn, Negative = will Not Churn
- Why it matters:
 - Accuracy alone can hide important trade-offs (especially with imbalanced classes when the interesting class has less frequency)
 - We must measure how the model succeeds and fails

Lecture 6

Thinking in Terms of Cost and Benefit

Prediction Outcome	Cost or Benefit Example
True Positive (TP)	Retained customer = revenue gain
False Positive (FP)	Wasted retention cost on loyal user
False Negative (FN)	Lost opportunity to retain
True Negative (TN)	No action taken, no cost

- Positive = will Churn, Negative = will Not Churn
- Good models help maximize return from each decision based on predictions.
- The true value of a model lies in the business impact of its predictions.

Lecture
6

Expected Value of a Model

- Expected Value (EV): The weighted average of all possible outcomes based on:
 - Probability
 - Associated payoff or cost
- EV formula for a model:

$$EV = \sum (\text{Probability of Outcome} \times \text{Value of Outcome})$$

- Even a model with lower accuracy may have higher expected value if it supports more profitable actions.

Lecture
6Business Question $\neq$ Data Science Question

- **Business question:** What decision should we make to improve outcomes?
- **Data question:** What patterns exist in the data?
- Gap example: Predicting churn (data science) $\neq$ Knowing who to retain profitably (business)
- **Solution:** Incorporate:
 - Business constraints (budget, capacity)
 - Action costs and outcomes
 - Risk tolerance

Lecture
6

Aligning Models with Business Objectives

- Models should support business strategies, not replace them.
- Begin with the business question, not the data science technique.
- Examples:
 - Business Question: “Who is likely to churn?” → Predictive Model:
Churn probability
 - Business Question: “Who should we target with retention offers?” →
Decision Model: Optimize ROI based on churn risk + offer cost

Lecture
6

Model Fidelity vs. Model Value

- **Model fidelity**: How accurately it fits the training data
- **Model value**: How much it helps drive better outcomes
- Sometimes a simpler, less accurate model is more useful:
 - Easier to explain
 - Faster to deploy
 - More robust over time

The best model is the one that makes the best decisions — not just predictions.”

Lecture 6

Precision, Recall, and F1 Score

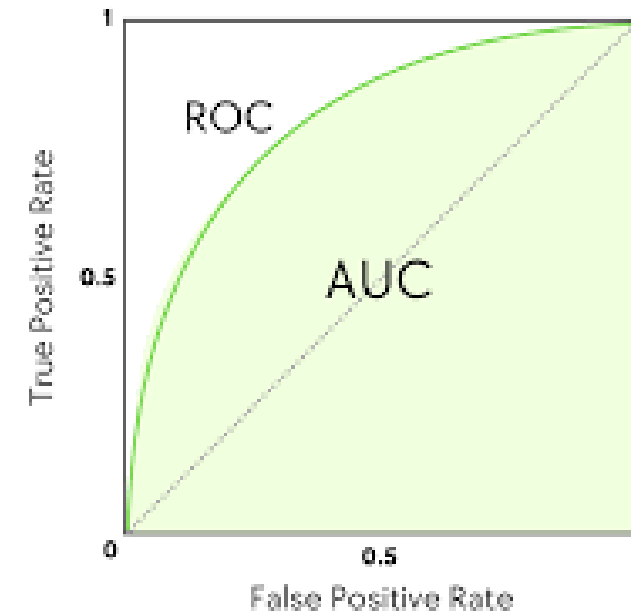
Metric	Formula	Meaning
Precision	$TP / (TP + FP)$	% of predicted positives that were correct
Recall	$TP / (TP + FN)$	% of actual positives identified
F1 Score	$2 \times (Precision \times Recall) / (Precision + Recall)$	Balance between precision and recall

- TP = True Positive, FP = False Positive, TN = True Negative, FN = False Negative
- Use Cases:
 - High precision: useful when cost of false positives is high
 - High recall: crucial when false negatives are very costly

Lecture
6

ROC Curve and AUC

- **ROC Curve (Receiver Operating Characteristic Curve)**
 - Plots: True Positive Rate vs. False Positive Rate at different thresholds
- **AUC (Area Under Curve):**
 - Ranges from 0.5 (random guessing) to 1.0 (perfect classifier)
 - Measures model's ability to distinguish classes across thresholds
 - AUC is useful when the ranking of predictions is more important than exact probabilities.



Lecture 6

Cost-Sensitive Evaluation

- Not all errors cost the same!

Type of Error	Business Example	Potential Cost
False Positive	Targeted non-churner with incentive	Marketing budget waste
False Negative	Missed churner	Lost revenue

- Solutions:
 - Use cost matrices to simulate real-world decisions
 - Adjust thresholds to reflect acceptable trade-offs

Lecture 6

False Positives vs. False Negatives – A Deliberate Trade-Off

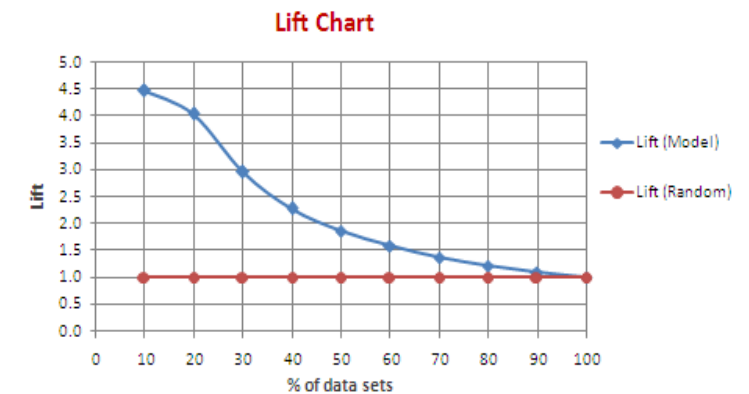
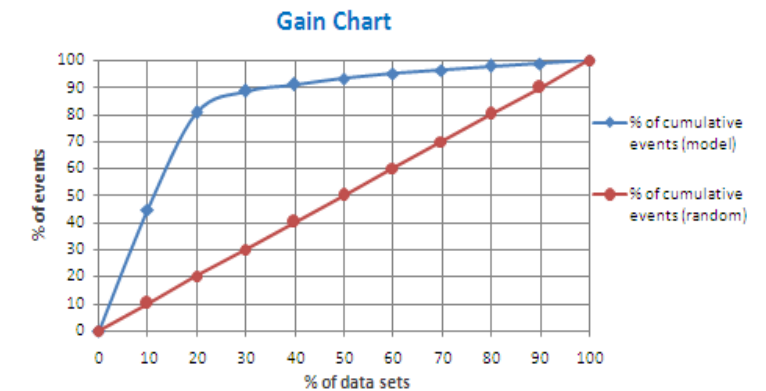
Strategy	Effect
Lowering threshold	Increases recall, more false positives
Raising threshold	Increases precision, more false negatives

- Example:
 - Fraud detection: prioritize low false negatives
 - Email marketing: prioritize high precision
- Tuning the threshold controls your risk appetite

Lecture 6

Lift and Gains Charts

- Used to assess how well the model helps prioritize actions:
- Lift chart: Compares model's ability vs. random targeting
- Gains chart: Shows cumulative % of positive cases captured as you move through ranked list
- Use Case:
 - Target top 10% of customers based on model scores
→ Gain = % of actual churners found



Lecture 6

Choosing the Right Metric for the Right Job

Business Problem	Suggested Focus
Churn Prediction	Recall, Lift
Credit Risk	AUC, Precision
Email Campaign	Precision, Gains
Fraud Detection	Recall, Cost-sensitive

- Key Message:
 - Don't chase generic benchmarks
 - Select metrics based on business priorities, costs, and decisions
- The best evaluation metric is the one that reflects the value of your actions.

Lecture 6

Comparing Models – It Depends on Your Goal

Model A	Model B
Higher accuracy	Higher recall
Better for common cases	Better at catching rare events
Slower but interpretable	Faster but complex

- You can't always say "Model A is better than Model B" without context.
- Key Question: Which model supports better decisions for this business use case?

Lecture
6When High Accuracy $\neq$ Better Decisions

- Predicting churn example:
 - Model with 95% accuracy (but poor recall on actual churners)
- **Better choice:** A model with 80% accuracy but catches twice as many actual churners
- **Insight:**
 - High accuracy can hide poor performance on the important class
 - Always assess impact on business outcomes, not just prediction metrics

Lecture
6

Adjusting Decision Thresholds

- Most models output probabilities, not direct decisions.
- You control where to draw the line:
 - Default = 0.5 (probability of class membership)
 - Adjust to favor recall or precision depending on costs
- Tuning thresholds helps optimize for **decision goals**, not generic metrics

Threshold ↑	Precision ↑	Recall ↓
Threshold ↓	Precision ↓	Recall ↑

Lecture 6

Business Rules vs. Data-Driven Models

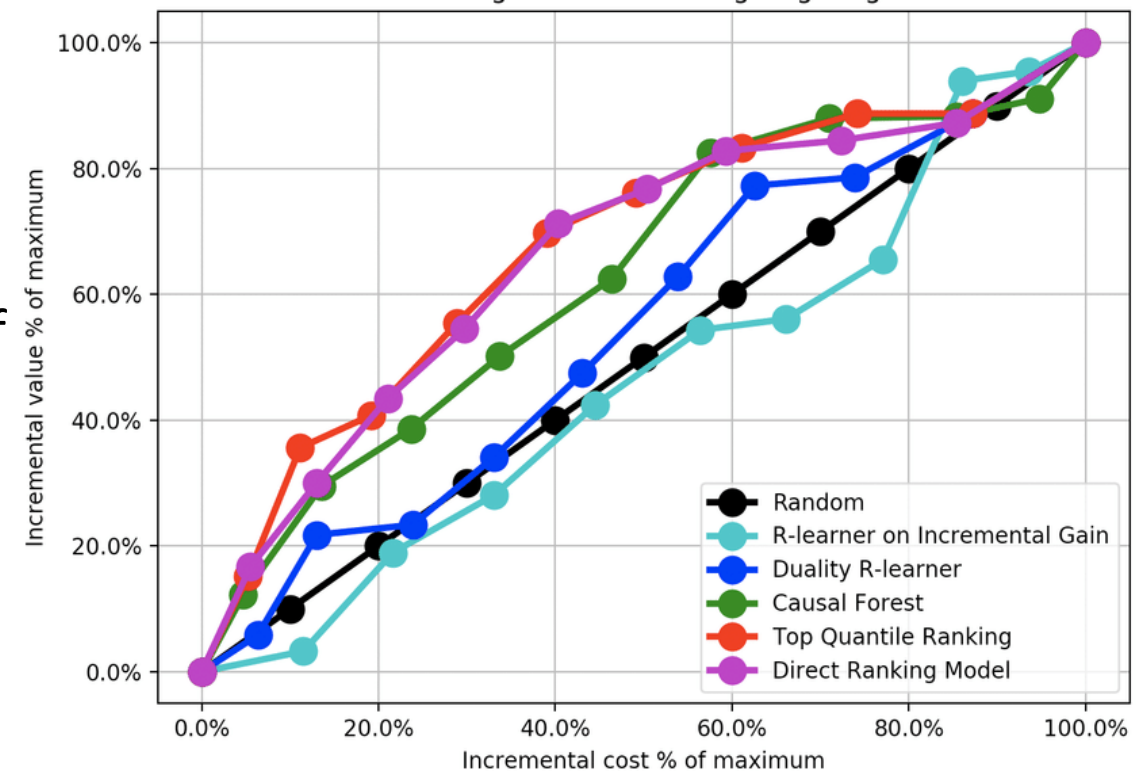
- Best Practice: Combine both → Use models to support, refine, or challenge business rules
- “A good model doesn’t replace human logic — it complements it.”

Business Rules	Machine Learning Models
Easy to understand	May be complex
Based on domain logic	Based on patterns in data
Static and rigid	Adaptive and flexible

Lecture 6

Cost Curves – Visualizing Trade-Offs

- Cost Curves plot model performance across different cost and class distributions
- Helps answer: “Which model is better if the cost of false positives changes?”
- Allow decision-makers to compare models under uncertainty
- Cost curves extend ROC analysis to include business-driven comparisons.



Lecture
6

Segment-Based Model Evaluation

- A model might work well overall, but poorly for key subgroups.
- **Implication:** Test performance within segments like:
 - Customer types
 - Regions
 - Product lines
- Helps ensure model is fair, effective, and interpretable across your business landscape

Segment	Precision	Recall
Young customers	0.80	0.60
Long-tenure users	0.45	0.30

Lecture 6

Managing Multiple Objectives

- A good model balances multiple competing business goals — not just one metric.
- Real-world decisions involve trade-offs:

Goal	Conflict
Maximize conversions	May increase false positives
Avoid false alarms	May miss true positives
Increase revenue	May increase risk

- Solution: Use multi-objective optimization:
 - Adjust thresholds
 - Build separate models for different outcomes
 - Use expected cost-benefit to guide decisions

Lecture
06

Sample Cases



**Lecture
6**

Churn Prediction with Cost Constraints

- **Scenario:**
 - Predict which customers will churn
 - Offer a discount to retain them
- **Decision Considerations:**
 - Offering everyone a discount is too costly
 - Offering too few means losing customers
- **A “Good” Model:**
 - Targets high-risk customers
 - Balances retention cost vs. lifetime value
 - Maximizes net revenue impact
 - Success is defined by ROI of action, not prediction precision alone

Lecture
6

Fraud Detection – Managing False Negatives

- Challenge:
 - Fraud is rare but very costly
 - Missing one fraud (false negative) = big loss
 - Investigating honest users (false positives) = cost too
- A “Good” Model:
 - Prioritizes recall (catch more fraud)
 - Minimizes false negatives, even at cost of more alerts
 - Can integrate with human review queue for verification
 - Value lies in early, accurate alerts, not model elegance

**Lecture
6**

Credit Scoring in Regulated Environments

- **Context:**
 - Banks use models to approve/decline loans
 - Must comply with regulations on fairness and interpretability
 - Simpler models may be preferred if they are auditable and meet business/legal needs
- **A “Good” Model:**
 - Is explainable (e.g., logistic regression, decision trees)
 - Avoids bias or adverse impact
 - Meets thresholds for acceptable risk (e.g., default probability < 10%)

Lecture 6

Dynamic Pricing – Balancing Revenue and Risk

- Use Case:
 - Predict demand to set prices dynamically
 - Must avoid pricing too high (loss of sales) or too low (lost margin)
- Success = Increase in average order value without hurting conversion
- A “Good” Model:
 - Learns from past customer sensitivity
 - Adjusts based on time, competition, and inventory
 - Allows for controlled testing (A/B) to validate gains

Lecture
6

Cross-Selling – Using Lift over Precision

- **Goal:** Suggest add-ons to increase average basket size
- Predictive precision may be low, especially in e-commerce
- But if uplift (Lift) over random is high, model is valuable
- **A “Good” Model:**
 - Ranks users/items based on likely response
 - Targets top percentile for offers
 - Focuses on conversion lift, not raw accuracy
 - “Lift” answers: How much this model is better than random at picking good targets?

Lecture
06

Summary Tools



Lecture
6

Operationalizing Model-Based Decisions

- Even the best model fails if it's not actionable
- **A “Good” Model:**
 - Integrates into workflow or system (CRM, payment gateway, marketing tools)
 - Is monitored regularly for performance decay
 - Triggers clear decisions: who to target, approve, flag, or ignore

Model value is realized only when it improves decisions at scale.

**Lecture
6****Interpretability and Trust**

- A “good” model is as understandable as it is effective
- A technically strong model may fail if stakeholders can’t understand or trust it.
- Why Interpretability Matters:
 - Executives must justify decisions (e.g., lending, pricing)
 - Regulators may require explainability
 - Front-line teams need confidence in system outputs

Model Type	Interpretability
Decision Tree	High
Logistic Regression	Medium
Neural Network	Low

Lecture
6

Fairness and Ethical Considerations

- Models can unintentionally discriminate if they reflect biased training data.
- **Key Risks:**
 - Disparate impact across age, gender, race
 - Proxy bias (e.g., ZIP code as income proxy)
 - Denying opportunity based on irrelevant factors
- **A “Good” Model:**
 - Is tested for fairness metrics (e.g., equal opportunity)
 - Avoids protected attributes where not legally or ethically appropriate
 - Encourages ethical reflection during design

Lecture
6

Adaptability to Change

- Data patterns shift over time:
 - New behaviors (e.g., post-pandemic)
 - Market changes (e.g., inflation, tech disruption)
 - Robustness over time = sustained business value
- A “Good” Model:
 - Performs well today and can be retrained tomorrow
 - Is monitored regularly for concept drift
 - Can be updated incrementally without full retraining

Lecture
6

Scalability and Ease of Deployment

- **A “Good” Model:**
 - Runs efficiently in production systems (CRM, web app, batch jobs)
 - Is maintainable by non-experts
 - Plays well with company infrastructure (e.g., APIs, cloud)
- Model pipelines that include:
 - Preprocessing
 - Prediction
 - Action automation
- A model’s true success is in execution, not experimentation

Lecture 6

Human-in-the-Loop Modeling

- Many decisions benefit from combining human and machine judgment.
- The best models augment decision makers, not replace them.
- Benefits:
 - Improves accuracy and accountability
 - Builds user confidence in predictions
 - Helps surface edge cases the model can't capture

Use Case	Model Role	Human Role
Loan Approval	Score applicant	Override if warranted
Fraud Detection	Flag transactions	Confirm with investigation
Hiring	Rank candidates	Evaluate final fit

Lecture 6

The Value of Simplicity

- Simple models often win in practice
- The most complex model isn't always the most useful — sometimes, clarity is power.”

Benefit	Why It Matters
Easier to implement	Faster deployment
Easier to understand	Better stakeholder buy-in
Easier to maintain	Less technical debt
Often good enough	High accuracy isn't always required

Lecture
6Reflection – Good Models
Enable Good Decisions

- The purpose of modeling is not to build the best model. It's to make the best decisions.
- **Ask yourself:**
 - Would a business team know what to do with the model's output?
 - What decision changes if the prediction changes?
 - Is the model actionable, not just accurate?
- **Takeaway:** Good models drive value through better, faster, or more scalable decisions.

Lecture
6

Model Evaluation Checklist

Use this checklist to assess whether your model is truly good:

- ☐ Does it align with a specific business objective?
- ☐ Are costs and benefits of predictions considered?
- ☐ Have you evaluated precision, recall, AUC, lift, or other task-appropriate metrics?
- ☐ Is the model understandable and explainable?
- ☐ Can it be deployed and maintained in your real-world system?
- ☐ Is it fair and unbiased toward sensitive populations?

**Lecture
6****Key Points**

A good model is one that supports better decisions — not just one with high accuracy.

Accuracy can be misleading; business context, costs, and outcomes matter more.

Evaluation metrics must align with business goals.

Use precision, recall, lift, or cost-benefit analysis depending on what the organization values (e.g., minimizing fraud, maximizing retention ROI).

Thresholds and trade-offs matter.

Adjusting decision thresholds helps optimize for different goals like minimizing false positives or capturing more true positives.

Interpretability, fairness, and deployability are as important as performance.

A simpler, understandable model may be more useful — especially in regulated, ethical, or high-stakes domains.

The best model improves decisions, not just predictions.

Models should be judged by how well they integrate into workflows and help real users make better choices.

For Discussion Panel and Next Lecture:

- Chapter 7 “Decision Analytic thinking I: What is a Good Model?”

**Lecture
6**



**For Discussion Panel
and Next Lecture**



Thank You