140.800: How to AI (for Public Health)

Learning Types, Evaluation, and Problem Formulation

Yiqun T. Chen
Email: yiqunc@jhu.edu
Schedule office hours via email

Departments of Biostatistics and Computer Science
Data Science & AI Initiative and Malone Center for Engineering in Health

Recent AI Breakthroughs in Healthcare

This Month's Headlines:

- Google's Med-Gemini: Multimodal AI for clinical reasoning
- OpenAI GPT-4V: Medical image interpretation capabilities
- FDA approvals: AI-powered diagnostic tools hitting clinics
- Epic integration: LLMs in electronic health records

Why This Matters:

- AI is transitioning from research to real clinical practice
- Healthcare organizations need AI-literate professionals (or others will continue to fill the void!)
- Understanding how these systems work is becoming essential

The Big Question: How do we harness this technology responsibly and effectively for public health?

Supervised Learning

What it is: Learn from labeled examples

Health example: X-ray images → pneumonia diagnosis

Key insight: We tell the machine the "right answer" during training

Unsupervised Learning

What it is: Find hidden patterns in unlabeled data

Health example: Group patients with similar symptoms

Key insight: Machine discovers patterns we might miss

Reinforcement Learning

What it is: Learn through trial and error with rewards

Health example: Optimize drug dosing through patient outcomes

Key insight: Like training a clinician through experience

Self-Supervised Learning - Simple Example

What it is: Learn from data's own structure (no external labels needed)

Simple Example - Fill in the Blank:

supervised

(0/1)

healthy

- Sentence: "The patient has [MASK] blood pressure"
- AI learns: "high", "low", "normal" are likely answers
- No doctor needed to label this training data!

Sample: The patient has high blood pressure

The patient has ___ blood pressure

Goal: predict "High" as opposed to "Low"/etc.

Why Self-Supervised Learning Matters

Key Benefits:

Learning true representations from data structure

Foundation for modern LLMs like GPT and BERT

Regression vs Classification (supervised learning)

Regression (Continuous Outcomes):

- What we predict: Blood pressure, length of stay, drug dosage

Classification (Categorical Outcomes):

- What we predict: Disease/no disease, risk category

Why Different Loss Functions?

The Universal ML Framework

Every supervised learning problem can be written as:

$$\boxed{Y = f(X) + \epsilon}$$

What Each Component Represents:

- Y: What we want to predict (outcome)
- X: What we have to predict with (features)
- f: The relationship we want to learn
- ϵ: Random error/noise we can't predict

Our Learning Strategy:

- We'll tackle these components in order: Y, X, f, ϵ
- Understanding each component helps us make better modeling decisions
- This framework applies to any domain, not just healthcare

Component 1: Y (Outcomes)

Different Types of Outcomes:

- Continuous: Blood pressure (mmHg), length of stay (days)
- Binary: Disease/no disease, alive/dead
- Categorical: Risk level (low/medium/high)
- Ordinal: Pain scale (1-10), stage of disease
- Open-ended: Free text responses, clinical notes
- Image outcomes: Segmentation maps, image quality scores

stable diffusion

Text prompt → Image

Brainstorm: Y (Outcomes)

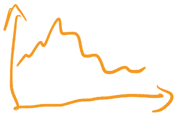
Exercise:

- What healthcare outcomes are you interested in predicting?
- What type of outcome variable would each be?
- How might the outcome type affect our modeling approach?

Component 2: X (Features)

Types of Features in Healthcare:

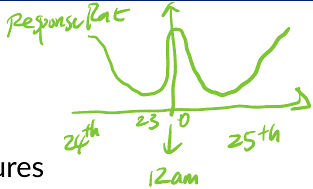
- Continuous: Lab values, vital signs, measurements
- Categorical: Gender, insurance type, hospital unit
- Binary: Presence/absence of conditions
- Ordinal: Severity scores, pain levels
- Text: Clinical notes, discharge summaries
- Images: X-rays, pathology slides, MRIs
- Time series: Continuous monitoring data



Brainstorm: X (Features)

Exercise:

- For your outcome of interest, what features might be predictive?
- What types of features would they be?
- What challenges might each feature type present?



Traditional Approach:

- Experts manually design relevant features
- Example: "Age > 65", "BMI category", "Number of medications"
- Works well with domain knowledge
- Limited by human creativity and time

Q: When is your prof most likely to respond to your email?

Input X: ① 12am - 12pm
② 0-24
③ ⋮

Feature Learning (AI-discovered)

Modern Approach:

- AI automatically learns relevant patterns
- Example: Neural network finds complex combinations
- Can discover non-obvious relationships
- Requires lots of data and computation

Healthcare Data Reality

In Practice:

- Both approaches have their place
- Traditional methods: interpretable, work with small datasets
- Modern methods: powerful, but need more data and compute
- Hybrid approaches often work best

Component 3: f (The Function)

Different Modeling Approaches by Complexity:

Simple/Interpretable Models

Linear regression, logistic regression

- Easy to understand and explain
- Work well with ~~smaller~~ datasets

some

(non-image, non-text)

Moderate Complexity Models

Random forests, gradient boosting (XGBoost)

- Balance between performance and interpretability
- Handle mixed data types well

High Complexity Models

Neural networks, deep learning models

- Can capture very complex patterns (esp. image, text datasets)
- Need lots of data and computational resources

Component 4: ϵ (Making Inferences)

The Error Term Represents:

- Random variation: Things we can't predict
- Measurement error: Imperfect data collection

Statistical Inference Questions:

- How confident are we in our predictions?
- Which features are most important? (SHAPLEY values)
- Will this generalize to new populations? (Distributional shift)
- What's the uncertainty around our estimates?

Why Statistical Inference Matters

In Healthcare:

Clinical decisions require uncertainty estimates

Regulatory approval requires statistical evidence

The Complete Framework

For Any Healthcare ML Problem, Ask: $Y = f(X) + \varepsilon$

1. What am I trying to predict? (Y - outcome type)
2. What data do I have? (X - feature types and quality)
3. What's the right model complexity? (f - simple vs complex)
4. How do I quantify uncertainty? (ϵ - inference and confidence)

where "ML/AI" usually happens

This Framework Scales:

- Works for simple logistic regression
- Works for complex neural networks
- Works for modern foundation models
- Always start with these four questions

Next Steps: We'll dive deeper into each component and see how modern AI tackles these challenges at scale

Training Data

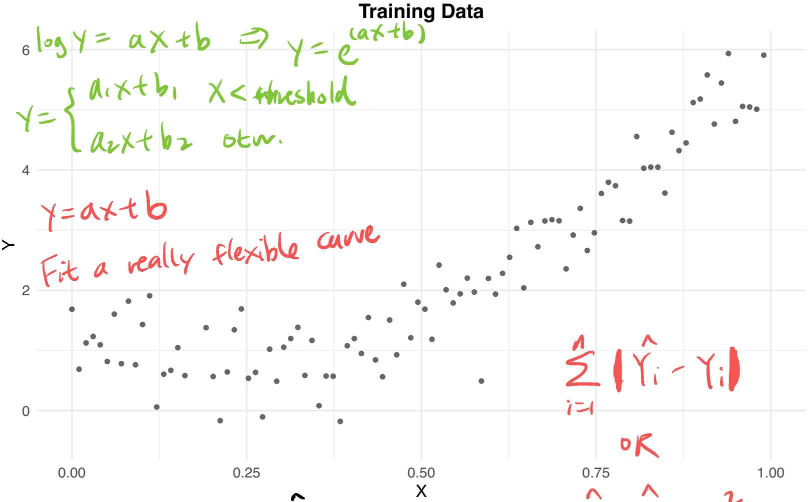
1. $Y = aX^2 + bX + c$ (Quadratic/Sq).

2. $\log Y = aX + b \Rightarrow Y = e^{(aX+b)}$

3. $Y = \begin{cases} a_1X + b_1 & X < \text{threshold} \\ a_2X + b_2 & \text{otw.} \end{cases}$

4. $Y = aX + b$

5. Fit a really flexible curve



$\sum_{i=1}^{n} |\hat{Y}_i - Y_i|$

OR

$\sum_{i=1}^{n} |\hat{Y}_i - Y_i|^2$

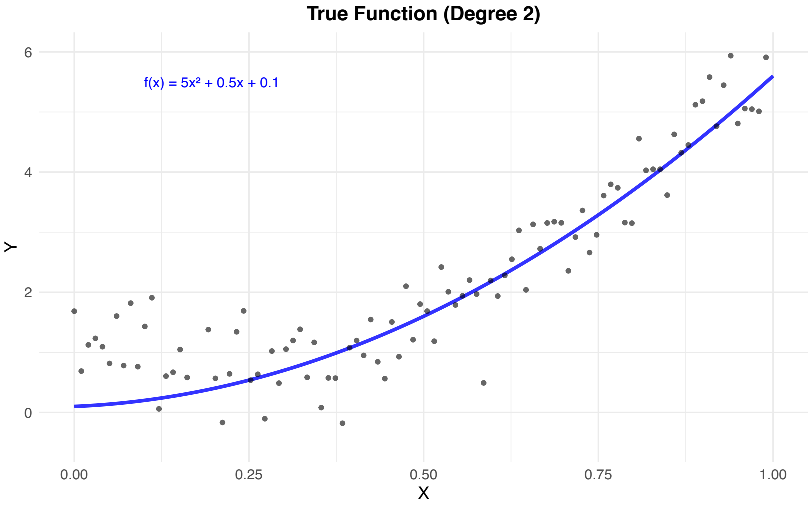
pred: $\hat{Y}$

truth: Y

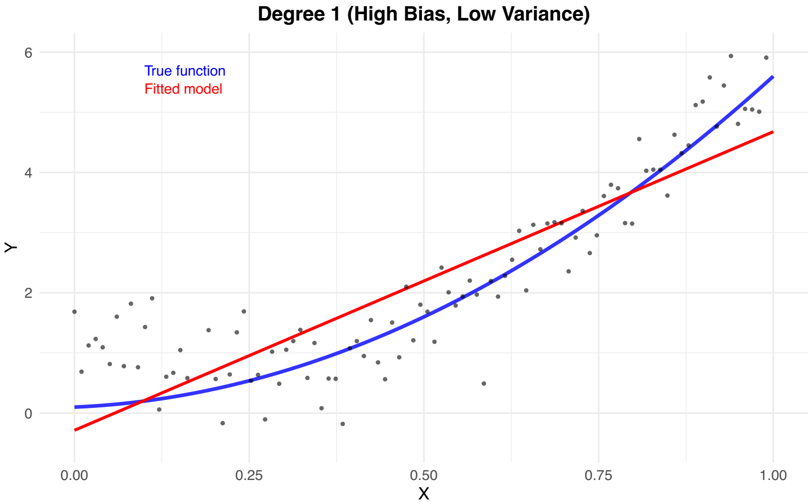
$\Rightarrow (Y - \hat{Y})$

AIC.

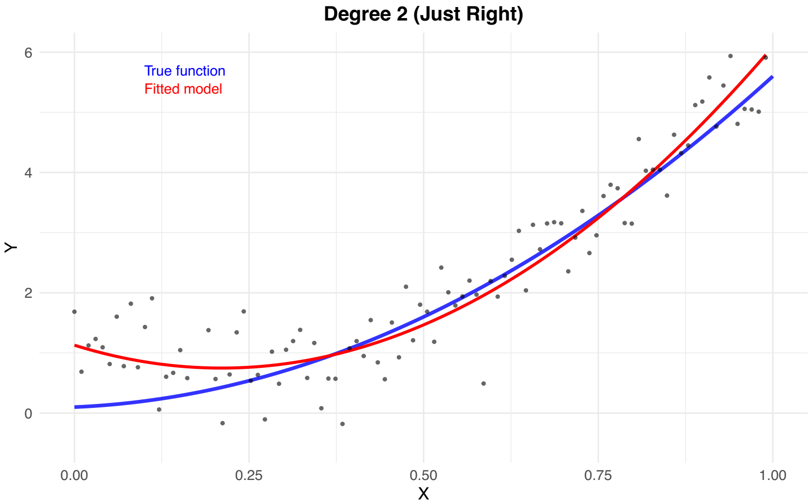
True Function (Unknown in Practice)



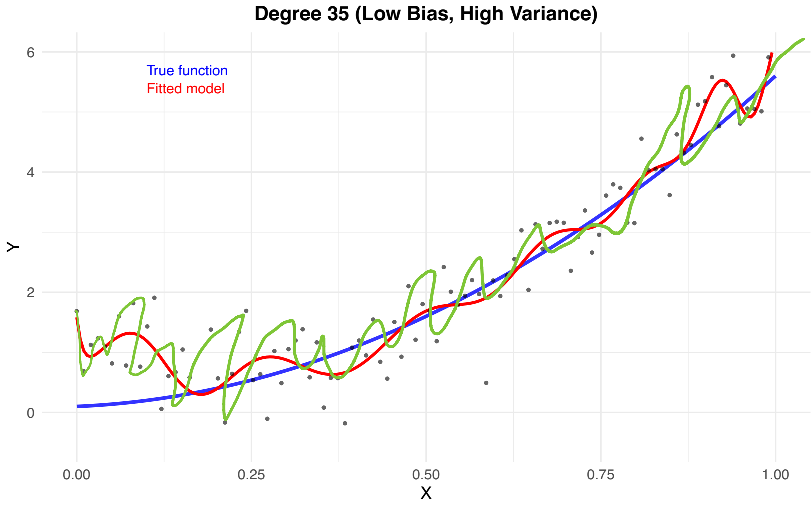
Polynomial Fitting Examples



Polynomial Fitting Examples



Polynomial Fitting Examples



The Bias-Variance Tradeoff

Two Sources of Prediction Error:

- Bias: How far off is our model on average?
- Variance: How much do predictions vary with different training sets?

Healthcare Example - Predicting Length of Stay:

- **High Bias/Low Variance**: Simple linear model (always predicts average)
- **Low Bias/High Variance**: Complex model (memorizes training data)
- **Sweet Spot**: Balanced model that generalizes well

The Importance of Test Sets

Why We Need Held-Out Data:

- Problem: Models can memorize training data perfectly
- Solution: Keep some data completely hidden during training
- Test on this data: Shows how well model generalizes to new patients

Data Splitting Strategy:

- **Training Set**: Learn model parameters
- **Validation Set**: Tune hyperparameters/model selections
- **Test Set**: Final evaluation

The Importance of Test Sets

Why We Need Held-Out Data:

- Problem: Models can memorize training data perfectly
- Solution: Keep some data completely hidden during training
- Test on this data: Shows how well model generalizes to new patients

Data Splitting Strategy:

- **Training Set**: Learn model parameters
- **Validation Set**: Tune hyperparameters/model selections
- **Test Set**: Final evaluation

Modern Dataset Considerations:

- Temporal splits: Train on 2020-2022, test on 2023
- Hospital splits: Train on Hospital A, test on Hospital B
- Participant splits: Ensure no patient data leakage across sets

Overcoming Bias-Variance Issues

We will talk more about how to overcome these challenges later

Coming up: Techniques for finding the right balance

Key Takeaways

Today's Foundation:

1. **Learning Types**: Supervised, unsupervised, reinforcement, self-supervised
2. **Supervised Learning**: Regression vs classification with different approaches
3. **Universal Framework**: $Y = f(X) + \epsilon$ - a way to think about any ML problem
4. **Components Matter**: Y, X, f, ϵ each require different considerations
5. **Model Complexity**: Bias-variance tradeoff is fundamental
6. **Test Sets**: Essential for honest evaluation

This Foundation Enables:

- Understanding how modern AI systems work
- Making good modeling choices for your problems
- Communicating effectively about AI projects
- Building robust, reliable healthcare AI systems

Key Takeaways

Today's Foundation:

1. **Learning Types**: Supervised, unsupervised, reinforcement, self-supervised
2. **Supervised Learning**: Regression vs classification with different approaches
3. **Universal Framework**: $Y = f(X) + \epsilon$ - a way to think about any ML problem
4. **Components Matter**: Y, X, f, ϵ each require different considerations
5. **Model Complexity**: Bias-variance tradeoff is fundamental
6. **Test Sets**: Essential for honest evaluation

This Foundation Enables:

- Understanding how modern AI systems work
- Making good modeling choices for your problems
- Communicating effectively about AI projects
- Building robust, reliable healthcare AI systems