How to give a research talk

Sanjoy Dasgupta

University of California, San Diego

Outline

1. The right mindset
   - What is the goal in giving a talk?
   - What should you consider to be your contribution?
2. Preparing the talk
   - How should you structure your talk?
   - Should you give proofs, and how?
   - What's the best way to practice your talk?
3. The talk and the aftermath
   - What can you do about your nervousness?
   - How do you manage questions during the talk?
   - How can you assess how the talk went?

1. The right mindset

What is the goal in giving a talk?

A. To entertain?
B. To impress?
C. To educate?

What is your contribution?

Your research has many components:

1. A study of the literature and perhaps a new way of organizing it
2. An evaluation of earlier technical results and a way of conceptualizing them
3. A variety of illustrative examples and experiments that have built up your intuition
4. Your novel mathematical result

What is your contribution?

Your research has many components:

1. A study of the literature and perhaps a new way of organizing it
2. An evaluation of earlier technical results and a way of conceptualizing them
3. A variety of illustrative examples and experiments that have built up your intuition
4. Your novel mathematical result

(1)–(3) are the crucial scaffolding that have made your result possible and may well be more interesting than the result itself.

Example: Organizing the literature

From a talk I recently attended.

Existing methods for fast nearest neighbor search can be grouped into four categories:

1. Tree structures
2. Hash tables
3. Quantization
4. Graph-based methods

Example: Organizing the literature

From a talk I recently attended.

> Existing methods for fast nearest neighbor search can be grouped into four categories:
>
> 1. Tree structures
> 2. Hash tables
> 3. Quantization
> 4. Graph-based methods

So much of what we do is **organizing knowledge**. It is worth sharing this.

2. Preparing the talk

How should you structure your talk?

Key constraint: Audience's limited attention span.

- If they feel they have learned something, they will be pleased and will expectantly wait another few minutes to see if there is more good stuff on the way
- Otherwise they will go to sleep

How should you structure your talk?

Key constraint: Audience's limited attention span.

- If they feel they have learned something, they will be pleased and will expectantly wait another few minutes to see if there is more good stuff on the way
- Otherwise they will go to sleep

To better retain their attention: Give your talk a **narrative flow**.

Narrative flow

It can help to have your talk progress like a story.

- You introduce some characters (definitions, concepts)
- You bring up the central problem
- You give a little history
- You attempt a solution
- It fails and you learn from this failure
- You try again, a different way, and this time it works
- You now wrap it all up in a happy result

Narrative flow

It can help to have your talk progress like a story.

- You introduce some characters (definitions, concepts)
- You bring up the central problem
- You give a little history
- You attempt a solution
- It fails and you learn from this failure
- You try again, a different way, and this time it works
- You now wrap it all up in a happy result

Benefits:

- Makes it easier for the audience to follow along
- Makes it easier for you to give the talk

Example: Consistency of k-nearest neighbor

Let's try to present a classical result.

Stone (1977): The k-nearest neighbor classifier is universally consistent in $\mathbb{R}^d$, provided $k/n \to 0$ and $k \to \infty$.

Example: Consistency of k-nearest neighbor

Let's try to present a classical result.

Stone (1977): The k-nearest neighbor classifier is universally consistent in $\mathbb{R}^d$, provided $k/n \to 0$ and $k \to \infty$.

Can start with a more elementary and informal statement, like

> The k-nearest neighbor classifier converges to the Bayes optimal decision rule, provided k grows with n.

This quickly gives people a (possibly very vague) idea of what is coming.

The Bayes-optimal classifier

Introduce key definitions step-by-step, carefully.

- Classification task: instance space $\mathcal{X}$, labels $\mathcal{Y} = \{0, 1\}$.

The Bayes-optimal classifier

Introduce key definitions step-by-step, carefully.

- Classification task: instance space $\mathcal{X}$, labels $\mathcal{Y} = \{0, 1\}$.
- Statistical learning framework: There is a distribution P over $\mathcal{X} \times \mathcal{Y}$ from which all instances (X, Y) are drawn.

The Bayes-optimal classifier

Introduce key definitions step-by-step, carefully.

- Classification task: instance space $\mathcal{X}$, labels $\mathcal{Y} = \{0, 1\}$.
- Statistical learning framework: There is a distribution P over $\mathcal{X} \times \mathcal{Y}$ from which all instances (X, Y) are drawn.
- Can decompose P into two pieces:
  - The distribution of X
  - The distribution of $Y|X$

  Write the second distribution as $\eta(x) = \Pr(Y = 1|X = x)$.

The Bayes-optimal classifier

Introduce key definitions step-by-step, carefully.

- Classification task: instance space $\mathcal{X}$, labels $\mathcal{Y} = \{0, 1\}$.
- Statistical learning framework: There is a distribution P over $\mathcal{X} \times \mathcal{Y}$ from which all instances (X, Y) are drawn.
- Can decompose P into two pieces:
    - The distribution of X
    - The distribution of $Y|X$

  Write the second distribution as $\eta(x) = \Pr(Y = 1|X = x)$.
- The Bayes-optimal classifier: $h^*(x) = 1$ if $\eta(x) \geq 1/2$ and 0 otherwise.

The Bayes-optimal classifier

Introduce key definitions step-by-step, carefully.

- Classification task: instance space $\mathcal{X}$, labels $\mathcal{Y} = \{0, 1\}$.
- Statistical learning framework: There is a distribution P over $\mathcal{X} \times \mathcal{Y}$ from which all instances (X, Y) are drawn.
- Can decompose P into two pieces:
  - The distribution of X
  - The distribution of $Y|X$

  Write the second distribution as $\eta(x) = \Pr(Y = 1|X = x)$.
- The Bayes-optimal classifier: $h^*(x) = 1$ if $\eta(x) \geq 1/2$ and 0 otherwise.
- Error rate is $R^* = \mathbb{E}_X \min(\eta(X), 1 - \eta(X))$.

The nearest neighbor classifier

Introduce remaining concepts: k-NN classifier, consistency.

The question of **consistency**: Does the k-NN classifier converge to the Bayes optimal rule as the number of training points goes to infinity?

Now we have the main characters:

- Bayes optimality
- k-NN classifier
- consistency

And we have introduced the central problem.

The nearest neighbor classifier

Introduce remaining concepts: k-NN classifier, consistency.

> The question of **consistency**: Does the k-NN classifier converge to the Bayes optimal rule as the number of training points goes to infinity?

Now we have the main characters:

- Bayes optimality
- k-NN classifier
- consistency

And we have introduced the central problem.

Next: result, or obstacles.

1-NN is not consistent

Example: $\mathcal{X} = \mathbb{R}$ and $\eta(x) = 3/4$ everywhere.

Every label is a coin flip with bias $3/4$.

1-NN is not consistent

Example: $\mathcal{X} = \mathbb{R}$ and $\eta(x) = 3/4$ everywhere.
Every label is a coin flip with bias $3/4$.

- Bayes error rate is $R^* = 1/4$ (always predict 1).

1-NN is not consistent

Example: $\mathcal{X} = \mathbb{R}$ and $\eta(x) = 3/4$ everywhere.
Every label is a coin flip with bias 3/4.

- Bayes error rate is $R^* = 1/4$ (always predict 1).
- 1-NN error rate: what is the probability that two coins of bias 3/4 disagree?

$$\frac{1}{4} \times \frac{3}{4} + \frac{3}{4} \times \frac{1}{4} = \frac{3}{8}$$

No matter how many training examples are seen!

1-NN is not consistent

Example: $\mathcal{X} = \mathbb{R}$ and $\eta(x) = 3/4$ everywhere.
Every label is a coin flip with bias $3/4$.

- Bayes error rate is $R^* = 1/4$ (always predict 1).
- 1-NN error rate: what is the probability that two coins of bias $3/4$ disagree?

$$\frac{1}{4} \times \frac{3}{4} + \frac{3}{4} \times \frac{1}{4} = \frac{3}{8}$$

  No matter how many training examples are seen!
- Likewise k-NN is not consistent for any fixed k

Consistency of k-nearest neighbor

Main theorem: If k grows with n such that $k/n \to 0$ and $k \to \infty$ then the error rate of the k-NN classifier converges to the Bayes error R^*.

Relation to prior work:

- Cover and Hart (1965): Consistency when η is continuous
- This work: Consistency without assumptions

Should you give proofs, and how?

Two good ways of presenting proofs

1. **The view from above**
   - Main theorem T rests on lemmas A and B. Have a slide with T, A, B.
   - Now prove lemma A.
   - Return to slide with T, A, B.
   - Now prove lemma B.
   - Return to slide with T, A, B. Done!

Should you give proofs, and how?

Two good ways of presenting proofs

1. **The view from above**
   - Main theorem T rests on lemmas A and B. Have a slide with T, A, B.
   - Now prove lemma A.
   - Return to slide with T, A, B.
   - Now prove lemma B.
   - Return to slide with T, A, B. Done!

2. **How it came about: the story of the process.**

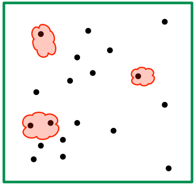
Consistency of k-nearest neighbor

Consistency of k-nearest neighbor

- Starting point: Prior work has already shown consistency for continuous η. But we have a general (measurable) η.

Consistency of k-nearest neighbor

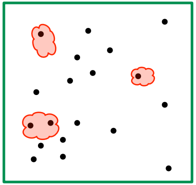
- Starting point: Prior work has already shown consistency for continuous η. But we have a general (measurable) η.
- Idea: Any measurable function can be arbitrarily well approximated by a continuous function (Lusin's theorem).
  Specifically: for any $\epsilon > 0$, there is a continuous function that agrees with η on all but ϵ fraction of $\mathcal{X}$. Does this do the trick?



Training points in the red region can cause trouble. What fraction of query points have one of these as their nearest neighbor?

Consistency of k-nearest neighbor

- Starting point: Prior work has already shown consistency for continuous η. But we have a general (measurable) η.
- Idea: Any measurable function can be arbitrarily well approximated by a continuous function (Lusin's theorem).
  Specifically: for any $\epsilon > 0$, there is a continuous function that agrees with η on all but ϵ fraction of $\mathcal{X}$. Does this do the trick?



Training points in the red region can cause trouble. What fraction of query points have one of these as their nearest neighbor?

- Geometric result: pick any finite set of points in $\mathbb{R}^d$. Then any one point is the NN of at most 5^d other points.

What's the best way to practice your talk?

1. Feedback from others
   Good for early feedback (e.g. "totally didn't understand the result")

2. Rehearsing on your own
   Good for ironing out wrinkles in the flow of ideas

3. The talk and the aftermath

What can you do about your nervousness?

1. Always: Sleep well
    - Hold off on coffee and alcohol the evening before
    - Go to bed early
    - Read a book before going to sleep
2. Possibly: Exercise?
3. If especially nervous: Practice your talk a lot

How do you manage questions during the talk?

1. Appreciate questions
2. Categories of questions
   - Simple clarifications
   - Mathematical details and proof intuitions
   - Open-ended questions

It is very important to finish on time! Ideally, finish early.

How do you know if the talk went well?

Sources of information:

1. Your gut feeling during the talk.

How do you know if the talk went well?

Sources of information:

1. Your gut feeling during the talk.
   - Probably unreliable

How do you know if the talk went well?

Sources of information:

1. Your gut feeling during the talk.
   - Probably unreliable
2. Questions during the talk.

How do you know if the talk went well?

Sources of information:

1. Your gut feeling during the talk.
   - Probably unreliable
2. Questions during the talk.
   - Lots of questions: very positive sign
   - No questions: neutral

How do you know if the talk went well?

Sources of information:

1. Your gut feeling during the talk.
   - Probably unreliable
2. Questions during the talk.
   - Lots of questions: very positive sign
   - No questions: neutral
3. Personal feedback after the talk.

How do you know if the talk went well?

Sources of information:

1. Your gut feeling during the talk.
   - Probably unreliable
2. Questions during the talk.
   - Lots of questions: very positive sign
   - No questions: neutral
3. Personal feedback after the talk.
   - "Good talk": positive
   - "Brilliant talk": very positive but never happens
   - "I learned a lot": very positive and you can do this!

Recap: Outline

1. The right mindset
   - What is the goal in giving a talk?
   - What should you consider to be your contribution?
2. Preparing the talk
   - How should you structure your talk?
   - Should you give proofs, and how?
   - What's the best way to practice your talk?
3. The talk and the aftermath
   - What can you do about your nervousness?
   - How do you manage questions during the talk?
   - How can you assess how the talk went?