

From "I'm Interested" to "I Have a Question"

Xiangliang Zhang

Computer Science and Engineering
University of Notre Dame

8/12/2026 @ KDD'26 Undergraduate Consortium

About Me



Profiling program and user behaviors for anomaly intrusion detection based on non-negative matrix factorization	46	2004
W Wang, X Guan, X Zhang 2004 43rd IEEE Conference on Decision and Control (CDC)(IEEE Cat. No ...		
Modeling program behaviors by hidden Markov models for intrusion detection	85	2004
W Wang, XH Guan, XL Zhang Proceedings of 2004 International Conference on Machine Learning and ...		
A novel intrusion detection method based on principle component analysis in computer security	85	2004
W Wang, X Guan, X Zhang International Symposium on Neural Networks, 657-662		



Summary of Primary Supervision

Ph.D.	Completed: 18; In Progress: 9
Masters	Completed: 15; In Progress: 0
Undergraduate	Completed: 20; In Progress: 1

Why?

How?

interest

question

Interest gives you energy.

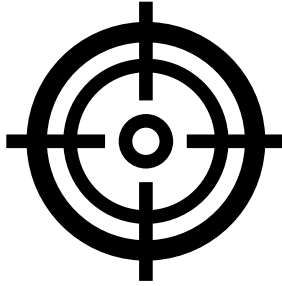
I am interested in

LLMs
Agentic AI
AI for Science
AI for Health
.....

BUT, still don't know where to start



A question gives you direction.



Why?

interest

How?

question

**What's different about testing an LLM and testing a student?*

**Do LLMs struggle with the same math questions I do?*

**Will an AI agent cheat when debugging code?*

These are not final research questions yet. But each one points somewhere: what to read, what to test, who to ask.....

Why?

How?

interest

question

**Template: I am interested in X,
but I noticed Y.
Under what conditions does Z happen?**

Broad interest

Evaluating LLMs

Noticed

We test LLMs with exams,
but students and LLMs fail
differently

Question

What is different about
testing an LLM and
testing a student?

Why?

How?

interest

question

**Template: I am interested in X,
but I noticed Y.
Under what conditions does Z happen?**

Broad interest

LLMs solving Math

Noticed

**Some problems feel easy
to me but hard for
models, or the reverse**

Question

**Do LLMs struggle with
the same math
questions students do?**

Why?

How?

interest

question

**Template: I am interested in X,
but I noticed Y.
Under what conditions does Z happen?**

Broad interest

AI coding agents

Noticed

**Agents can use tools in
unexpected ways**

Question

**Will an AI agent cheat
when debugging code?**

Why?

How?

interest

question

**Template: I am interested in X,
but I noticed Y.**

Under what conditions does Z happen?

Broad interest

AI for Health

Noticed

**Models can be accurate
overall but unreliable for
small subgroups**

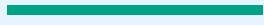
Question

**When does subgroup-
aware evaluation change
model selection?**

In your project, what is the 'but I noticed' sentence?

Why?

interest



How?

question

**Template: I am interested in X,
but I noticed Y.
Under what conditions does Z happen?**

Don't have the 'but I noticed' sentence?

You need to

- Find a supervisor and talk to her/him
- Get your hands on the work

Run the code.

Read the data.

Reproduce one result.

Try one small change.

Look for what surprises you.

What questions I asked in these works?

Profiling program and user behaviors for anomaly intrusion detection based on non-negative matrix factorization

W Wang, X Guan, X Zhang

2004 43rd IEEE Conference on Decision and Control (CDC)(IEEE Cat. No ...

46

2004

Modeling program behaviors by hidden Markov models for intrusion detection

W Wang, XH Guan, XL Zhang

Proceedings of 2004 International Conference on Machine Learning and ...

85

2004

A novel intrusion detection method based on principle component analysis in computer security

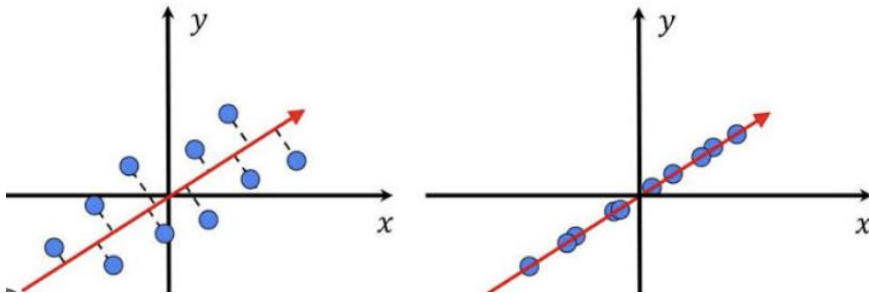
W Wang, X Guan, X Zhang

International Symposium on Neural Networks, 657-662

85

2004

Principal Component Analysis (PCA)



Find PCs
Projection

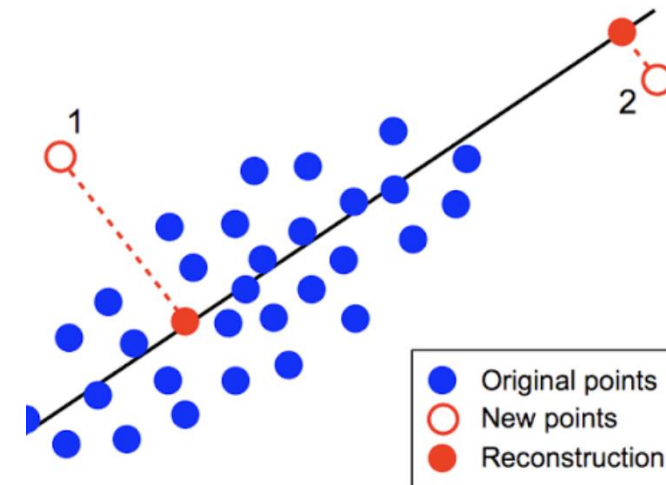
Reconstruction

Noticed

This point 1 is reconstructed with a big error

Question

What if the badly reconstructed point is not a noise?



A good question or a bad question?

01 Specific

It names a setting, data, behavior, or failure mode.

02 Answerable next

It suggests one concrete next test.

03 Contestable

A reasonable person could expect a different result.

04 Meaningful

The answer would change what someone does or believes.

Research is not homework with a longer deadline

In homework, the problem is already shaped. In research, **shaping the problem** is part of the work.

Homework

Clear prompt
Known tools
Expected answer
Fast feedback

Research

Fuzzy prompt
Unclear tools
Unknown answer
Ambiguous feedback

What you notice in homework can become a research question.

How to train yourself to ask good questions?

Suggestion 1: Read papers to find moves, not just facts

What problem did they choose?	What did they ignore?
What assumption makes the method work?	Where might it fail?
What comparison makes the claim believable?	What comparison is missing?

How to train yourself to ask good questions?

Suggestion 2: Talk to your supervisor

Before the meeting

What I tried

What changed

Where I am stuck

Bring three short things

A result or observation

A sharper question

Two possible next steps

A strong undergraduate does not hide confusion; they make it inspectable.

How to train yourself to ask good questions?

Suggestion 3: Track your progress

Weak progress

"I tried many models."

Strong progress

"Model A fails when labels are sparse, but only for group B."

The second sentence creates a next experiment.

01

Read

Find one choice or assumption.

02

Try

Run one small test.

03

Notice

Name what changed.

04

Ask

Turn the surprise into a question.

Repeat this loop and the question will sharpen.

Which one do you want to hear?

Weaker impression

"I am very interested in machine learning and willing to work hard."

Stronger impression

"I read one paper on retrieval. I did not understand the evaluation. I want to test when retrieval actually helps."

Thank you for your attention!

Do not wait until you feel like a researcher.
Practice asking questions that make the next step clearer.

Questions?