

MATH + AI + FINANCE

How Math Powers AI in Everyday Finance

From Your Morning Coffee to Your Credit Score

Prof. Jörg Osterrieder

45 MINUTES

5 IDEAS

1 MACHINE

The Minds Behind the Machine

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

Eight mathematicians. 2000 years. One AI.



Grassmann
Vectors, 1844



Pascal
Probability, 1654



Bayes
Inference, 1763



Newton
Calculus, 1666



Leibniz
Notation, 1684



Cauchy
Gradient, 1847



Shannon
Info Theory, 1948



Hinton
Nobel Prize, 2024

The Central Question

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future



“The same machine that wins gold at the Math Olympiad cannot count to three.”

— The Paradox of Modern AI

The Central Question

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future



“The same machine that wins gold at the Math Olympiad cannot count to three.”

— The Paradox of Modern AI

This talk traces **five mathematical ideas** — invented for gambling, war, telephones, accounting, and plague — that accidentally built the AI in your pocket.

Five answers. One machine. 2000 years of math.

Idea 1: Linear Algebra

Opening

Algebra

THREAD 1

~100 BCE

Probability

Calculus

Info & Optim

Convergence

Your Future

Idea 1: Linear Algebra

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

THREAD 1

~100 BCE Jiuzhang Suanshu — “Nine Chapters”

Chinese merchants used bamboo counting rods for **inventory, taxation, and land surveying**. Rows and columns of rods = Gaussian elimination, **2,000 years before Gauss**.

~100 BCE



Bamboo counting rods

↓ Modern notation

$$\begin{pmatrix} 3 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 2 & 3 \end{pmatrix}$$

“Nine Chapters on the Mathematical Art”

~100 BCE

Idea 1: Linear Algebra

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

THREAD 1

~100 BCE Jiuzhang Suanshu — “Nine Chapters”

Chinese merchants used bamboo counting rods for **inventory, taxation, and land surveying**. Rows and columns of rods = Gaussian elimination, **2,000 years before Gauss**.

The Silk Road is opening. Trade demands systematic accounting. Someone invents matrix operations to manage the books.

~100 BCE



Bamboo counting rods

↓ Modern notation

$$\begin{pmatrix} 3 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 2 & 3 \end{pmatrix}$$

“Nine Chapters on the Mathematical Art”

~100 BCE

The Schoolteacher Nobody Read

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

THREAD 1

1844



Hermann Grassmann

1809–1877

Schoolteacher, Stettin, Prussia.

The Schoolteacher Nobody Read

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

THREAD 1

1844



Hermann Grassmann

1809–1877

Schoolteacher, Stettin, Prussia.

1844: Publishes *Ausdehnungslehre* — invents vector spaces, linear independence, dimension. **Nobody reads it.** Gives up math, becomes a Sanskrit scholar.

The Schoolteacher Nobody Read

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

THREAD 1

1844



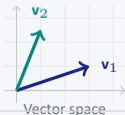
Hermann Grassmann

1809–1877

Schoolteacher, Stettin, Prussia.

1844: Publishes *Ausdehnungslehre* — invents vector spaces, linear independence, dimension. **Nobody reads it.** Gives up math, becomes a Sanskrit scholar.

Finance: Today every bank customer is a **vector** — income, age, credit history, spending. Grassmann invented this.



Matrix \times Vector: The Core Operation

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

THREAD 1

$$\mathbf{y} = \mathbf{W}\mathbf{x} + \mathbf{b}$$

Matrix \times vector + bias
Every neural network layer

Matrix \times Vector: The Core Operation

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

THREAD 1

$$\mathbf{y} = \mathbf{W}\mathbf{x} + \mathbf{b}$$

Matrix \times vector + bias
Every neural network layer

Example:

$$\mathbf{x} = \begin{pmatrix} 50,000 \\ 720 \end{pmatrix} \text{ (income, credit)}$$

$$\mathbf{W} = \begin{pmatrix} 0.3 & 0.7 \\ 0.5 & 0.5 \end{pmatrix}$$

Output: risk + approval scores

Matrix \times Vector: The Core Operation

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

THREAD 1

$$\mathbf{y} = \mathbf{W}\mathbf{x} + \mathbf{b}$$

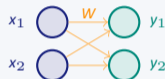
Matrix \times vector + bias
Every neural network layer

Example:

$$\mathbf{x} = \begin{pmatrix} 50,000 \\ 720 \end{pmatrix} \text{ (income, credit)}$$

$$\mathbf{W} = \begin{pmatrix} 0.3 & 0.7 \\ 0.5 & 0.5 \end{pmatrix}$$

Output: risk + approval scores



1.8T

MULTIPLICATIONS

PER TOKEN

Cayley (1858) formalized this *while practising law*. His notation executes **1.8 trillion times** per GPT-4 word.

Words Are Geometry

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

THREAD 1 → LLM

$$\vec{\text{King}} - \vec{\text{Man}} + \vec{\text{Woman}} \approx \vec{\text{Queen}}$$

Words Are Geometry

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

THREAD 1 → LLM

$$\vec{\text{King}} - \vec{\text{Man}} + \vec{\text{Woman}} \approx \vec{\text{Queen}}$$



Word2Vec (2013): Every word \rightarrow 300+ numbers.

GPT-4: 12,288 dimensions.

Meaning *is* direction in space.

Words Are Geometry

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

THREAD 1 → LLM

$$\vec{\text{King}} - \vec{\text{Man}} + \vec{\text{Woman}} \approx \vec{\text{Queen}}$$



Word2Vec (2013): Every word \rightarrow 300+ numbers.

GPT-4: 12,288 dimensions.

Meaning *is* direction in space.

The chain:

Accounting rods (100 BCE)

↓ Grassmann's vectors (1844)

↓ Cayley's matrices (1858)

↓ Word embeddings (2013)

↓ **Every LLM today**

Application: Your Bank Uses This Math

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

FINANCE APPLICATION

\$32B

ANNUAL FRAUD

Banks have **50 milliseconds** to decide: legit or fraud? That's three formulas from your math class.

Application: Your Bank Uses This Math

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

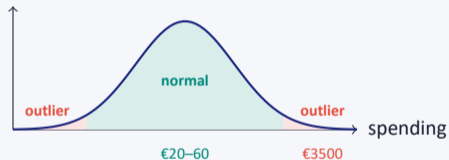
Your Future

FINANCE APPLICATION

\$32B

ANNUAL FRAUD

Banks have **50 milliseconds** to decide: legit or fraud?
That's three formulas from your math class.



Transactions far from center → flagged

Idea 2: Probability

- Opening
- Algebra
- Probability**
- Calculus
- Info & Optim
- Convergence
- Your Future

THREAD 2

1654

Idea 2: Probability

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

THREAD 2

1654 Chevalier de Méré loses money gambling

Complains to Pascal: “The math says I should win!” Pascal writes to Fermat. Their letters invent probability.

1654



Blaise Pascal, c. 1663

The question:

Two players, interrupted game.

How to split the pot *fairly*?

Idea 2: Probability

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

THREAD 2

1654 Chevalier de Méré loses money gambling

Complains to Pascal: “The math says I should win!” Pascal writes to Fermat. Their letters invent probability.

Their “Problem of Points” — splitting a pot fairly — became **insurance pricing**. Every premium traces back to two Frenchmen arguing about dice.

1654



Blaise Pascal, c. 1663

The question:
Two players, interrupted game.

How to split the pot *fairly*?

Bayes' Theorem: Updating Beliefs

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

THREAD 2

1763



**Thomas
Bayes**

1702–1761

Presbyterian minister.
Published posthumously.

Bayes' Theorem: Updating Beliefs

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

THREAD 2

1763



Thomas

Bayes

1702–1761

Presbyterian minister.
Published posthumously.

Finance: Foundation
of every **real-time risk
assessment** in banking.

$$P(H | E) = \frac{P(E | H) P(H)}{P(E)}$$

Start with prior → see
evidence → update belief

Bayes' Theorem: Updating Beliefs

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

THREAD 2

1763



**Thomas
Bayes**

1702–1761

Presbyterian minister.
Published posthumously.

Finance: Foundation
of every **real-time risk
assessment** in banking.

School analogy: Friend usually on time (95%). Today 20 min late + traffic jam on their route. You *update*: “stuck in traffic,” not “overslept.”

$$P(H | E) = \frac{P(E | H) P(H)}{P(E)}$$

Start with prior → see
evidence → update belief



When 99% Accurate Means 95% Wrong

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

WORKED EXAMPLE

$$P(\text{Fraud}) = 0.001$$

$$P(\text{Flag}|\text{Fraud}) = 0.99$$

$$P(\text{Flag}|\neg\text{Fraud}) = 0.02$$

When 99% Accurate Means 95% Wrong

WORKED EXAMPLE

$$P(\text{Fraud}) = 0.001$$

$$P(\text{Flag}|\text{Fraud}) = 0.99$$

$$P(\text{Flag}|\neg\text{Fraud}) = 0.02$$

$$P(\text{Fraud}|\text{Flag}) = \frac{0.99 \times 0.001}{0.99 \times 0.001 + 0.02 \times 0.999} \approx \mathbf{4.7\%}$$

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

When 99% Accurate Means 95% Wrong

WORKED EXAMPLE

$$P(\text{Fraud}) = 0.001$$

$$P(\text{Flag}|\text{Fraud}) = 0.99$$

$$P(\text{Flag}|\neg\text{Fraud}) = 0.02$$

$$P(\text{Fraud}|\text{Flag}) = \frac{0.99 \times 0.001}{0.99 \times 0.001 + 0.02 \times 0.999} \approx \mathbf{4.7\%}$$

5 real
95 false

47×

BETTER THAN

RANDOM

Only 5% of flags are real fraud — but that's **47× better than random**. Missing €10K fraud \gg one verification call.

AI doesn't need to be perfect. It needs to be better than nothing.

From Probabilities to Predictions

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

THREAD 2 → LLM

$$\text{softmax}(z_i) = \frac{e^{z_i}}{\sum_j e^{z_j}}$$

Any numbers →
valid probabilities

From Probabilities to Predictions

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

THREAD 2 → LLM

$$\text{softmax}(z_i) = \frac{e^{z_i}}{\sum_j e^{z_j}}$$

Any numbers →
valid probabilities

$[2.0, 1.0, 0.1] \rightarrow [0.66, 0.24, 0.10]$



Non-negative, sums to 1. Kolmogorov's axioms (1933) satisfied.

From Probabilities to Predictions

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

THREAD 2 → LLM

$$\text{softmax}(z_i) = \frac{e^{z_i}}{\sum_j e^{z_j}}$$

Any numbers →
valid probabilities

$[2.0, 1.0, 0.1] \rightarrow [0.66, 0.24, 0.10]$



Non-negative, sums to 1. Kolmogorov's axioms (1933) satisfied.

ChatGPT runs softmax over 50,000+ words every token. Temperature controls randomness.

The chain:

Gambling (1654) → Bayes (1763) → Kolmogorov (1933) → Softmax → **every AI word**

Your Turn: Spot the Fraud

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

Audience Vote

Raise your hand if you think it's fraud.

Alex, 17

€899 guitar online, 10 PM
Never bought instruments
Home IP address

Tomoko, 45

50 identical €50 gift cards
3:47 AM, Osaka address

Priya, 28

€3,200 wire to new UK account
Never sent to UK before

Your Turn: Spot the Fraud

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

Audience Vote

Raise your hand if you think it's fraud.

Alex, 17

€899 guitar online, 10 PM

Never bought instruments

Home IP address

LEGIT

Tomoko, 45

50 identical €50 gift cards

3:47 AM, Osaka address

FRAUD

Priya, 28

€3,200 wire to new UK account

Never sent to UK before

Your Turn: Spot the Fraud

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

Audience Vote

Raise your hand if you think it's fraud.

Alex, 17

€899 guitar online, 10 PM
Never bought instruments
Home IP address

LEGIT

Tomoko, 45

50 identical €50 gift cards
3:47 AM, Osaka address

FRAUD

Priya, 28

€3,200 wire to new UK account
Never sent to UK before

AMBIGUOUS

Could be tuition OR fraud

[BOT] *The AI's job is harder than it looks. Not every case has a clear answer.*

Idea 3: Calculus

- Opening
- Algebra
- Probability
- Calculus**
- Info & Optim
- Convergence
- Your Future

THREAD 3

1666 / 1684

Idea 3: Calculus

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

THREAD 3

Newton (1666): Calculus during the plague. Hid it.

Leibniz (1684): Published first. Notation: $\frac{dy}{dx}$.

Scandal: Newton rigged the Royal Society investigation. We use **Leibniz's notation** today.

1666 / 1684



Newton

VS



Leibniz



Idea 3: Calculus

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

THREAD 3

Newton (1666): Calculus during the plague. Hid it.

Leibniz (1684): Published first. Notation: $\frac{dy}{dx}$.

Scandal: Newton rigged the Royal Society investigation. We use **Leibniz's notation** today.

Finance: "Derivatives" are literally named after calculus. Options pricing, risk sensitivity, portfolio optimization.

1666 / 1684



Newton

VS



Leibniz



How AI Learns: The Blindfolded Hiker

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

THREAD 3

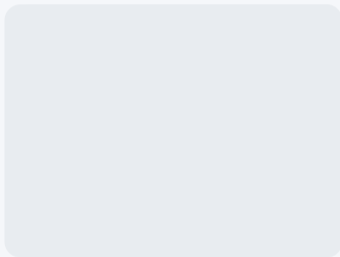
1847 → 2024



A.-L.

Cauchy

1789–1857



How AI Learns: The Blindfolded Hiker

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

THREAD 3

1847 → 2024



A.-L.
Cauchy
1789–1857

1847: “Step in the direction that reduces error. Repeat.”

Finance: Same algorithm calibrates every risk model.

$$\theta_{t+1} = \theta_t - \eta \nabla L(\theta_t)$$

Parameters = current
– rate \times gradient

How AI Learns: The Blindfolded Hiker

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

THREAD 3



A.-L.
Cauchy
1789–1857

1847: “Step in the direction that reduces error. Repeat.”

Finance: Same algorithm calibrates every risk model.

1847 → 2024

$$\theta_{t+1} = \theta_t - \eta \nabla L(\theta_t)$$

Parameters = current
– rate \times gradient

Chain rule enables backprop: $\frac{\partial L}{\partial w} = \frac{\partial L}{\partial y} \cdot \frac{\partial y}{\partial w}$

Leibniz’s 1684 notation powers every AI training run.



Application: Your Financial Report Card

FINANCE APPLICATION

$$\text{Score} = \sum w_i x_i$$

Not all factors are equal

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

Application: Your Financial Report Card

FINANCE APPLICATION

$$\text{Score} = \sum w_i x_i$$

Not all factors are equal

School grades:

Final Exam: 40%

Homework: 30%

Participation: 20%

Quiz: 10%

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

Application: Your Financial Report Card

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

FINANCE APPLICATION

$$\text{Score} = \sum w_i x_i$$

Not all factors are equal

School grades:

Final Exam: 40%

Homework: 30%

Participation: 20%

Quiz: 10%

FICO credit score:

Payment History: 35%

Amounts Owed: 30%

Length of History: 15%

New Credit: 10%

Credit Mix: 10%

Same formula, different weights, different stakes.



From Leibniz to Nobel Prize

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

THREAD 3 → LLM



Geoffrey Hinton

b. 1947

1986: Backpropagation paper in *Nature*.

From Leibniz to Nobel Prize

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

THREAD 3 → LLM



Geoffrey Hinton

b. 1947

1986: Backpropagation paper in *Nature*.

Decades of rejection. “Neural nets are dead.” Persists.

2024: Nobel Prize in Physics.

From Leibniz to Nobel Prize

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

THREAD 3 → LLM



Geoffrey Hinton

b. 1947

1986: Backpropagation paper in *Nature*.

Decades of rejection. “Neural nets are dead.” Persists.

2024: Nobel Prize in Physics.

The chain: Newton (1666) → Leibniz (1684) → Cauchy (1847) → Hinton (1986) → **every LLM.**

Chain rule → 338 years → Nobel Prize.



Your Turn: USE IT or SKIP IT?

Ethics Exercise

Should banks be **allowed** to use this data for credit decisions?

Income

Social Media

Zip Code

Phone Brand

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

Your Turn: USE IT or SKIP IT?

Ethics Exercise

Should banks be **allowed** to use this data for credit decisions?

Income

USE

Directly relevant

Social Media

SKIP

Privacy violation

Zip Code

Phone Brand

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

Your Turn: USE IT or SKIP IT?

- Opening
- Algebra
- Probability
- Calculus**
- Info & Optim
- Convergence
- Your Future

Ethics Exercise

Should banks be **allowed** to use this data for credit decisions?

Income

USE

Directly relevant

Social Media

SKIP

Privacy violation

Zip Code

DEBATE

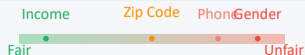
Proxy for race

Phone Brand

DISCRIM.

Punishes poverty

[BOT] *The hard part isn't the math — it's deciding what goes into the math.*



Idea 4: Information Theory

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

THREAD 4

1948



**Claude
Shannon**

1916–2001

Bell Labs. One paper,
one field.

Idea 4: Information Theory

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

THREAD 4



**Claude
Shannon**

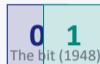
1916–2001

Bell Labs. One paper, one field.

Invents the **bit**, entropy, channel capacity.

Also: juggling unicyclist, flame-throwing trumpet.

1948



Idea 4: Information Theory

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

THREAD 4



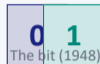
Claude
Shannon
1916–2001

Bell Labs. One paper, one field.

Invents the **bit**, entropy, channel capacity.

Also: juggling unicyclist, flame-throwing trumpet.

1948



$$H(P, Q) = - \sum P(x) \log Q(x)$$

How **surprised** is the model?

Designed for telephones (1948).
Now trains every LLM (2024).

Finance: prices every encrypted banking transaction.

1948 Architecture = 2024 Architecture

THREAD 4 → LLM

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

1948 Architecture = 2024 Architecture

Opening

Algebra

Probability

Calculus

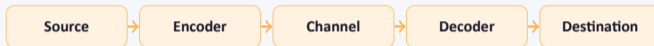
Info & Optim

Convergence

Your Future

THREAD 4 → LLM

SHANNON 1948

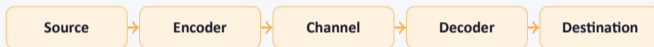


1948 Architecture = 2024 Architecture

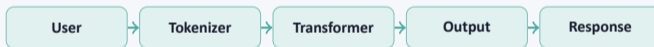
- Opening
- Algebra
- Probability
- Calculus
- Info & Optim**
- Convergence
- Your Future

THREAD 4 → LLM

SHANNON 1948



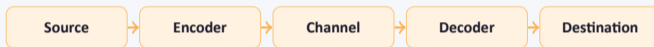
CHATGPT 2024



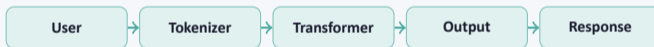
1948 Architecture = 2024 Architecture

THREAD 4 → LLM

SHANNON 1948



CHATGPT 2024



Same architecture. Neural networks replace each component.

Shannon's 1951 experiment: predict English letter by letter. Entropy ≈ 1.1 bits/char. 74 years later, LLMs do the same at word level. **Shannon would recognize ChatGPT.**

Idea 5: The Cost of Intelligence

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

THREAD 5: OPTIMIZATION

\$100M+

GPT-4

\$30M

LLAMA 3

\$5.6M

DEEPSEEK R1

Idea 5: The Cost of Intelligence

- Opening
- Algebra
- Probability
- Calculus
- Info & Optim**
- Convergence
- Your Future

THREAD 5: OPTIMIZATION

\$100M+

GPT-4

\$30M

LLAMA 3

\$5.6M

DEEPSEEK R1



Adam optimizer (Kingma, 2014, 200K+ citations) made this possible. Without it: **centuries**. With it: **weeks**.

Finance: AI training costs are the new capex. Understanding optimization math = understanding the economics of AI.

Why AI Can't Count to Three

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

THE FAILURE

Q: How many R's in "strawberry"?

ChatGPT: "Two"

WRONG (it's three)

Why AI Can't Count to Three

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

THE FAILURE

Q: How many R's in "strawberry"?

ChatGPT: "Two"

WRONG (it's three)

str

aw

berry

3 tokens, not 10 letters

Why? ChatGPT doesn't see letters. It sees **tokens**:

"strawberry" → ["str", "aw", "berry"]

The letter R is split across tokens. The AI *cannot see* individual characters. BPE trades letter accuracy for sentence efficiency.

Why AI Can't Count to Three

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

THE FAILURE

Q: How many R's in "strawberry"?

ChatGPT: "Two"

WRONG (it's three)

str

aw

berry

3 tokens, not 10 letters

Why? ChatGPT doesn't see letters. It sees **tokens**:

"strawberry" → ["str", "aw", "berry"]

The letter R is split across tokens. The AI *cannot see* individual characters. BPE trades letter accuracy for sentence efficiency.

The same system scoring gold at IMO cannot count letters. These aren't bugs — they're **clues about architecture**.

Probability \neq Truth

THE LIMIT

Xu et al., 2024: Current AI architectures will sometimes generate false statements. **No amount of training data can fully eliminate this.**

These limits are **mathematical**, not engineering.

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

Probability \neq Truth

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

THE LIMIT

Xu et al., 2024: Current AI architectures will sometimes generate false statements. **No amount of training data can fully eliminate this.**

These limits are **mathematical**, not engineering.

$$P(\text{all correct}) = p^n$$

$$\text{At 95\%/fact: } 0.95^{50} = 0.08$$



Probability \neq Truth

THE LIMIT

Xu et al., 2024: Current AI architectures will sometimes generate false statements. **No amount of training data can fully eliminate this.**

These limits are **mathematical**, not engineering.

$$P(\text{all correct}) = p^n$$

At 95%/fact: $0.95^{50} = 0.08$

2023: Lawyer Steven Schwartz cited 6 fake court cases generated by ChatGPT. None existed. He submitted them to a federal judge.

Small errors compound. A “95% accurate” model producing 50 facts \rightarrow 92% chance of at least one error.



2000 Years → One Model

Opening

Algebra

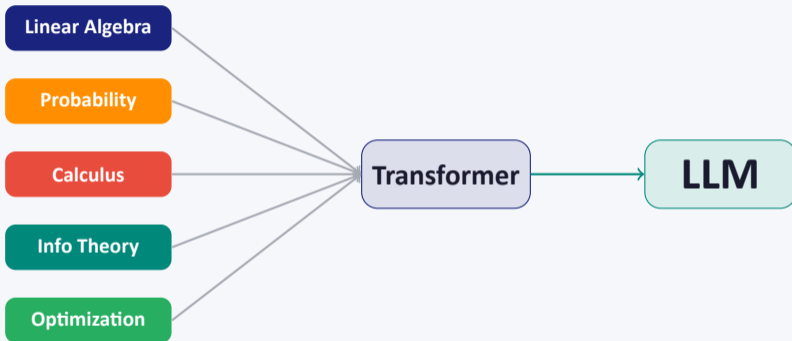
Probability

Calculus

Info & Optim

Convergence

Your Future



2000 Years → One Model

Opening

Algebra

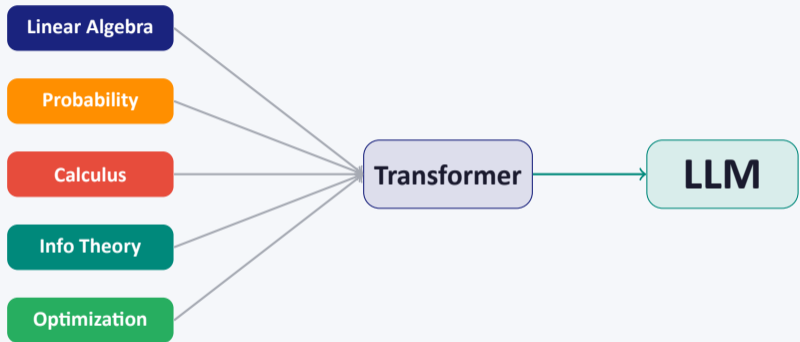
Probability

Calculus

Info & Optim

Convergence

Your Future



*“None of them knew they were building AI.
All of them made it possible.”*

The 2017 Paper That Changed Everything

CONVERGENCE

“Attention Is All You Need”

Vaswani et al., Google Brain, 2017

8 authors. One was a **20-year-old intern**.

Most-cited AI paper in history.

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

The 2017 Paper That Changed Everything

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

CONVERGENCE

“Attention Is All You Need”

Vaswani et al., Google Brain, 2017

8 authors. One was a **20-year-old intern**.

Most-cited AI paper in history.

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right) V$$

The 2017 Paper That Changed Everything

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

CONVERGENCE

"Attention Is All You Need"

Vaswani et al., Google Brain, 2017

8 authors. One was a **20-year-old intern**.

Most-cited AI paper in history.

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right) V$$

The cat sat down



Attention weights

All five ideas in one formula:

QK^T = matrix multiplication

softmax = probability

∇ = gradient descent trains it

Loss = cross-entropy

Adam = optimization

This is the **same cosine similarity** from slide 6.

When Math Discriminates

ETHICS

2019: Apple Card gave a husband 20× the credit limit of his wife — despite her having a higher score.

2024: Goldman Sachs fined \$89M.

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

When Math Discriminates

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

ETHICS

2019: Apple Card gave a husband 20× the credit limit of his wife — despite her having a higher score.

2024: Goldman Sachs fined \$89M.

Husband: 20×

Wife: 1×

Same credit score. Different limit.

The algorithm never used “gender.” But **proxy variables** — spending patterns, employment type — correlate with gender.

Removing a protected variable doesn't make an algorithm fair. Other variables “leak” the same information.

This is exactly the zip code problem from slide 17.

When Math Discriminates

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

ETHICS

2019: Apple Card gave a husband 20× the credit limit of his wife — despite her having a higher score.

2024: Goldman Sachs fined \$89M.

Husband: 20×

Wife: 1×

Same credit score. Different limit.

The algorithm never used “gender.” But **proxy variables** — spending patterns, employment type — correlate with gender.

Removing a protected variable doesn’t make an algorithm fair. Other variables “leak” the same information.

This is exactly the zip code problem from slide 17.

“Fair and useful aren’t always the same thing.”

Where This Math Works Right Now

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

INDUSTRY

200K

EMPLOYEES

\$17B

TECH BUDGET

2,000+

AI MODELS

Where This Math Works Right Now

INDUSTRY

200K

EMPLOYEES

\$17B

TECH BUDGET

2,000+

AI MODELS

Bayes → Fraud

Weighted Avg ← Credit

Cosine Sim ← Contracts

Sigmoid → Risk

Every formula from this talk runs inside JPMorgan **right now**:

Bayes → fraud Weighted avg → credit Cosine sim → contracts Sigmoid → risk

They use **RAG** (Retrieval-Augmented Generation) to reduce hallucinations. These are not theoretical formulas. This is **job-market math**.

BankBot's Journey

Opening

Algebra

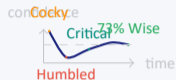
Probability

Calculus

Info & Optim

Convergence

Your Future



[BOT] *I analyzed your breakfast. You are 94% human. (Overconfident)*

BankBot's Journey

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future



[BOT] *I analyzed your breakfast. You are 94% human. (Overconfident)*

[BOT] *FRAUD DETECTED! ...Oh wait, those are suspicious flowers. (Humbled)*

BankBot's Journey

Opening

Algebra

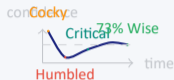
Probability

Calculus

Info & Optim

Convergence

Your Future



[BOT] *I analyzed your breakfast. You are 94% human. (Overconfident)*

[BOT] *FRAUD DETECTED! ...Oh wait, those are suspicious flowers. (Humbled)*

[BOT] *I could use social media data. Doesn't mean I should. (Critical)*

BankBot's Journey

Opening

Algebra

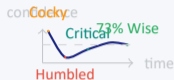
Probability

Calculus

Info & Optim

Convergence

Your Future



[BOT] *I analyzed your breakfast. You are 94% human. (Overconfident)*

[BOT] *FRAUD DETECTED! ...Oh wait, those are suspicious flowers. (Humbled)*

[BOT] *I could use social media data. Doesn't mean I should. (Critical)*

“I am 73% confident. But I defer to the human.”

Stage 8: Wise

All the Math, One Flow

- Opening
- Algebra
- Probability
- Calculus
- Info & Optim
- Convergence**
- Your Future



All the Math, One Flow



Every formula in this talk, working together in **50 milliseconds**. From raw transaction to approve/flag/block. The same pipeline powers credit decisions, insurance pricing, and loan approvals.

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

Screenshot This Slide

Opening

Algebra

Probability

Calculus

Info & Optim

Convergence

Your Future

Bayes' Theorem

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

Update beliefs with evidence

Cross-Entropy

$$H(P, Q) = - \sum P(x) \log Q(x)$$

How surprised is the model?

Weighted Average

$$\bar{x}_w = \sum w_j x_j$$

Not all factors are equal

Gradient Descent

$$\theta_{t+1} = \theta_t - \eta \nabla L$$

Step downhill, repeat

Matrix \times Vector

$$\mathbf{y} = W\mathbf{x} + \mathbf{b}$$

Every neural network layer

Cosine Similarity

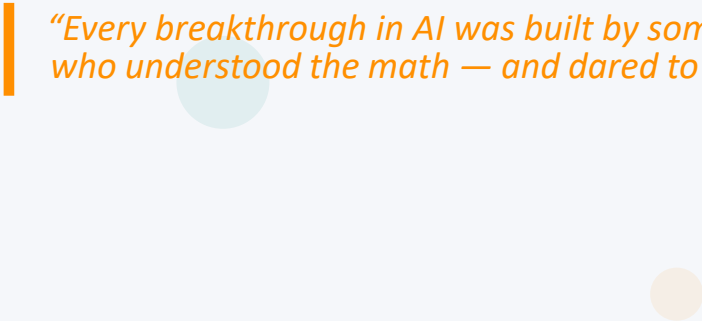
$$\cos \theta = \frac{\vec{a} \cdot \vec{b}}{|\vec{a}| |\vec{b}|}$$

How similar are two things?

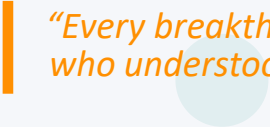
Softmax

$$\text{softmax}(z_i) = e^{z_i} / \sum e^{z_j}$$

Numbers \rightarrow probabilities



“Every breakthrough in AI was built by someone who understood the math — and dared to ask “what if?””



“Every breakthrough in AI was built by someone who understood the math — and dared to ask “what if?””

Nightingale used data to save soldiers. Lovelace saw that math is computation. Wald looked where no one else looked. Shannon predicted ChatGPT 74 years early.

Math has always needed diverse thinkers. The next breakthrough might be yours.

Website: digital-ai-finance.github.io/mathematics-for-ai