

# The Art of Discovery: Scientific Method, Logic, Mathematics, and the Role of Data Science

Rushikesh Joshi

Department of Computer Science & Engineering  
IIT Bombay, Mumbai, India

# Research everywhere

If you look around, everywhere you will find research that has been previously done

starting from our cloths, the stitches of our cloths, the tables and chairs, the pens and pencils, computers, cell phones, networks, writing boards, building itself, electricity and lights, refrigerators, fans, water supplied, internet, food packets, plastic wrappers and bags, spectacles, ornaments, perfumes, medicines, packing materials, food, commercial softwares, electronics and IC chips, cars, oils, glass and artifacts, sports equipments ..

... name it and you may even find phd thesis or research articles connecting to them!

# What was research yesterday may not be research today

Newton's laws

Kirchoff's laws

Ohm's law

The laws of thermodynamics

Value of gravitational constants

Principles of electricity generation

Laws of optics

Steam engine

.... Today we use them, and do further research

# A Quest from the Known into the Unknown

**Ideas:** Set theory, the Biological Evolution Tree, Big Bang Theory, String Theory

**Theorems:** Pythagoras Theorem, Pumping Lemma

**Laws:** Ohm's Law, Newton's laws, interchangeability of mass and energy

**Techniques:** Proof by induction, making plastic from crude oil, making steel

**Algorithm:** Primes in P, Dijkstra's shortest Path Algorithm

**Structural organization:** Structure of an Atom, Structure of a DNA, Periodic table,

**Dynamic behavior:** Celestial motions-N body problem, Electromagnetic induction, IC Engine, PN junction, electronic gates, flip-flops

**Natural Processes:** formation of a black hole, nuclear fusion in the Sun, fertilization of life, birth process, process of digestion, body metabolism processes,

# Science vs. mathematics and Logic

Science is about the physical universe

Maths can be used to model ideas in science

Maths can be done independently

Maths is an axiomatic system, which gives rise to an artificial universe

We map real life behaviors onto the artificial universe called Maths

Only then can we analyse, understand, and utilize that knowledge

Logic is part of Mathematics, and its foundation too

# Scientific method of Enquiry

Observe, Note

Hypothesize, Fit a mathematical model

Experiment and Validate with variations - repeatability

Concretize theories

# Mathematical method of Discovery

Axiomatize

Create a little universe, Observe properties

Hypothesize, Conjecture

Prove

# Examples of early research : Astronomical Observations

Planetary motions

Trajectories of the Sun

Solar and Lunar Eclipses

Observatory yantras

Observations of Nakshatras and finding directions, Navigation using Sky positions

Panchang Calculations

# Examples of early research in observing mathematical properties

Decimal number system, the Zero, negative numbers

Pythagoras Theorem and geometry

Tower of Brahma and other games

Chandas as Poetic meters in terms of Binary system

Algebra, Geometry, Combinatorics, Calculus

# Examples of early research in Engineering

The Wheel and carts, chariots

Construction of houses and huge temples

Rust-free iron/steel

Making of Gold, silver and copper Coins

Medicinal processes of purification

# Examples of early research in medicine

Herbs

Metals

Processes of purification and conversion

Nadi measures

Anupan

Pathya

structure of the body, yogic techniques

# Examples of early research in languages and music

Words, roots, inflections

Combinations of words, Grammar or Vyakarana

Poetic meters

Rasa Shastra

Forms of Prose and Poetry

Dictionaries, Thesaurus and other resources

Notes, Raga Formulation and structure, different forms of music

Musical instruments and their characteristics

# Example Key Ideas that lead to tremendous technology impact on the humanity

Mechanics

Steam Engine

Electricity and Electronics

Communication on wires and wireless frequencies

Material science

Organic and inorganic chemistry

# A few Modern Challenges

Non-conventional sources of Energy

Waste-management and recycling

Rapid construction and engineering

Satellite technologies

Efficient engineering- power transmission, power conversion and utilization

Compact storage for information

Fail safe systems

# The Role of Statistics

Works of Prof. Mahalanobis and his students in statistics, planning commission of India and the surveys- for resource allocation and planning

Random variables, Distributions and statistical properties

Queuing theory, Markov models, stochastic processes

Use statistics as models to explain phenomena

It is used everywhere today- from planning to data science

# Indication of a Truly Progressive Research and Development

Facilities become extremely affordable

Happiness increases

Competition reduces

Poverty is eradicated

Free or affordable availability of resources, including education

Less conflicts

# Data Science Foundations

Collect Data

Analyze Data

Find patterns and behaviors

Model fitting

Predict the unknown

Run simulations

- For example, given maths and language subject marks, can science marks be predicted?
- Given energy consumption trends, can the needs be predicted?
- Given generation of waste, can pollution and AQI be predicted?
- Given production of grains and crops, can availability and prices be predicted?
- Statistics and data science can be applied to any discipline that generates data

# Artificial Intelligence

Early AI involved Knowledge Graphs, Search techniques, Logic

Modern AI involves model training using gigantic data sources

Then the trained models are used to perform tasks

There are a huge collection of new AI algorithms and techniques

It touches upon every way of perception- images, sounds, speech and music, videos, smells, touch, sensory AI, ontologies and languages, conversations and generation

Success of AI over past few decades, more surprisingly the LLMs and generative AI has created lots of new opportunities

# Key Concepts to Watch

There are 60 apples and 40 oranges shipped to a customer

Photographs are taken by a computer of the customer, and the computer AI is counting the numbers to check if the items are as per the order

The computer converges on 50 items as apples and 50 as oranges.

How do you measure the success of this decision?

# Recall and Precision

Let's take the case of Apples

Out of 50 apples predicted, 30 happen to be actually apples

So we got only 50% of the apples correctly identified out of 60 apples (50% recall)

But we got 30 out of 50 apples predictions correct (60% precision for our prediction)

We can calculate these figures for oranges, and for the whole set of predictions

# False Positive, True Positive, False Negative, True Negative

See Apple as Apple

See Apple as not Apple

See not Apple as Apple

See Not apple as Not Apple



# Subtleties of Logic

“If you solve a problem, I will give you a chocolate”; you didn’t solve it, but I still give you a chocolate. Am I violating the rule?

“If you don’t eat this, I will not talk to you”; you eat this, but I am not going to talk to you;

Is this violating the rule?!

# Subtleties of Logic

Example:

People found to have followed some bad practice spoke a particular language, so all people speaking that language are the followers of that bad practice

Is this correct ?

# Subtleties of Logic

Example:

A teacher says “All students in my class know the Pythagoras theorem”. Someone finds one student who cannot mention what the theorem is.

Does that disprove the teacher?

# Subtleties of Logic

Example:

A teacher says “I have one student in my class who can solve any trigonometric problem in Loni’s book of problems.”, another teacher finds one student who cannot do that, and then claims that the first teacher is wrong.

Is this correct?

# Subtleties of Logic

Example:

A teacher says “I have one student in my class who can solve any trigonometric problem in Loni’s book of problems.”, another teacher drops by on a random day and finds no student who could solve one particular problem that he picked, and then he claims that the first teacher is wrong.

Is this correct?

# Subtleties of Logic

Example:

A teacher says “I have one student in my class who can solve any trigonometric problem in Loni’s book of problems.”, another teacher drops by on a random day and finds that no student who could solve one particular problem that he picked, and then he claims that the first teacher is wrong.

Is this correct? –well that student in the claim could have been absent- so the second teacher needs to do the same test on some other day to cover the whole class!

# Golden Rules of Logic to Remember in Research

Implication is not equivalence

Correlation or co-occurrence is not causality

Absence of Proof is not Proof of Absence

Example does not prove Generalization

Counterexample disproves a “for all”

Generalization