

The background of the slide is white with a subtle pattern of light blue dots. Overlaid on this are stylized circuit lines in two colors: orange and blue. The orange lines are primarily located on the left side, forming a dense, vertical structure that resembles a circuit board or a data stream. The blue lines are more sparse and are located on the right side, forming a similar but less dense structure. The text is centered in the upper half of the slide.

# WELCOME TO INFORMATION RETRIEVAL SYSTEMS

SAPAN H MANKAD

# WARNING

ALL THE GUIDELINES AND INSTRUCTIONS TO BE DISCUSSED IN THIS PRESENTATION ARE REAL AND THEY MUST BE FOLLOWED BY EACH OF YOU IN ANY CIRCUMSTANCES. VIOLATION TO ANY OF THE GUIDELINES WILL BE STRICTLY DEALT WITH.

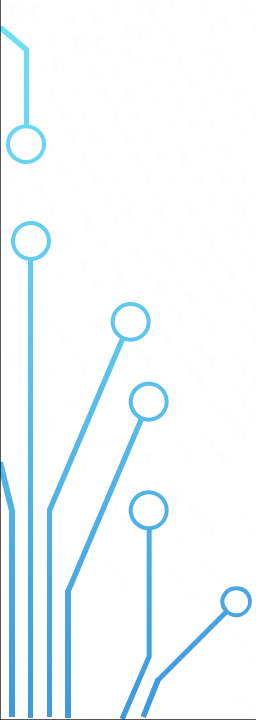
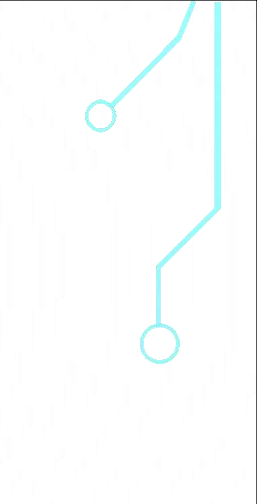
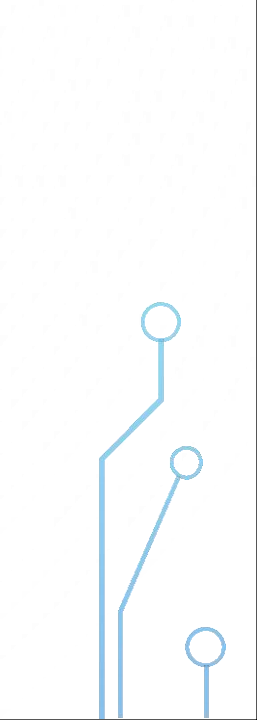
# PREREQUISITES

- A bit of programming
- Probability
- Statistics
- Linear Algebra  
(Matrices/Vectors)
- Keen Interest
- Passion

- Khan Academy
- YouTube videos
- NPTEL/MOOC



# *COURSE LOGISTICS*

- Term papers
  - Routine Homeworks
  - Assignments
  - Class Test
  - Sessional Examination
  - Semester End Examination
- 
- 
- 

# CLASSROOM ETIQUETTES

- *Classroom sessions are meant for only interested students.*
- *Arrive to the classroom in time.*
- *Plus Points*
- *Minus Points*

# LABORATORY ETIQUETTES

- *Maintain a lab notebook.*
- *Prepare a manual after completion of experiment.*
- *Show it to me regularly.*

# MODES OF COMMUNICATION

- Course Website
- Course Blog
- Online Learning Management System – Moodle
- Online Discussion Forum – Piazza
- Google Groups
- Email

# FORMATION OF STUDY GROUPS

- *Apply Machine Learning / Data Mining techniques to the domain of your interest or investigate some research in the same.*



# FORMATION OF STUDY GROUPS

- *Machine Learning*
- *Deep Learning*
- *Web Mining*
- *Text Mining*
- *Statistical Data Analysis*

# APPLICATION AREAS/DOMAINS

- *Speech, Music and Audio Processing*
- *Image Processing/Computer Vision*
- *Signal Processing*
- *E Commerce*
- *Social Network Analysis*
- *Recommender Systems*

# TOOLS

- *Python*
- *R*
- *Weka*
- *MATLAB*
- *Octave*

The image features a light blue background with a subtle grid pattern. In the corners, there are decorative elements resembling circuit board traces or neural network connections, consisting of thin blue lines and small circles.

# *BACKGROUND CHECK*

The image features a light blue background with decorative circuit-like lines in the corners. These lines are composed of small circles connected by straight lines, resembling a stylized electronic circuit. The lines are located in the top-left, top-right, bottom-left, and bottom-right corners, framing the central text.

*Today's databases are data rich,  
but information poor.*

# A BIG QUESTION...

- *From where do you find some information?*

Google

# WHAT IS INFORMATION RETRIEVAL?

- Information Retrieval (IR) is finding material (usually documents) of an unstructured nature (usually text) that satisfies an information need from within large collections (usually stored on computers).
- Examples
  - Web search
  - Email search
  - Local Search in your computer

# COURSE STRUCTURE

Architecture of IRS

Document  
Representation

Retrieval  
Models

Search and  
Filtering

Machine  
Learning in IR

Multimedia and  
Multilingual IR



The image features a light blue background with a subtle pattern of concentric circles. In the four corners, there are decorative elements resembling circuit board traces or neural network connections, consisting of thin blue lines and small circles.

# INTERESTING APPLICATIONS

# REASONING BASED QA

John moved to the bedroom.  
Mary grabbed the football there.  
Sandra journeyed to the bedroom.  
Sandra went back to the hallway.  
Mary moved to the garden.  
Mary journeyed to the office.

Where is the football?

# VISUAL QA



- Is this truck considered “vintage”?
- Does the road look new?
- What kind of tree is behind the truck?



How many benches are shown?