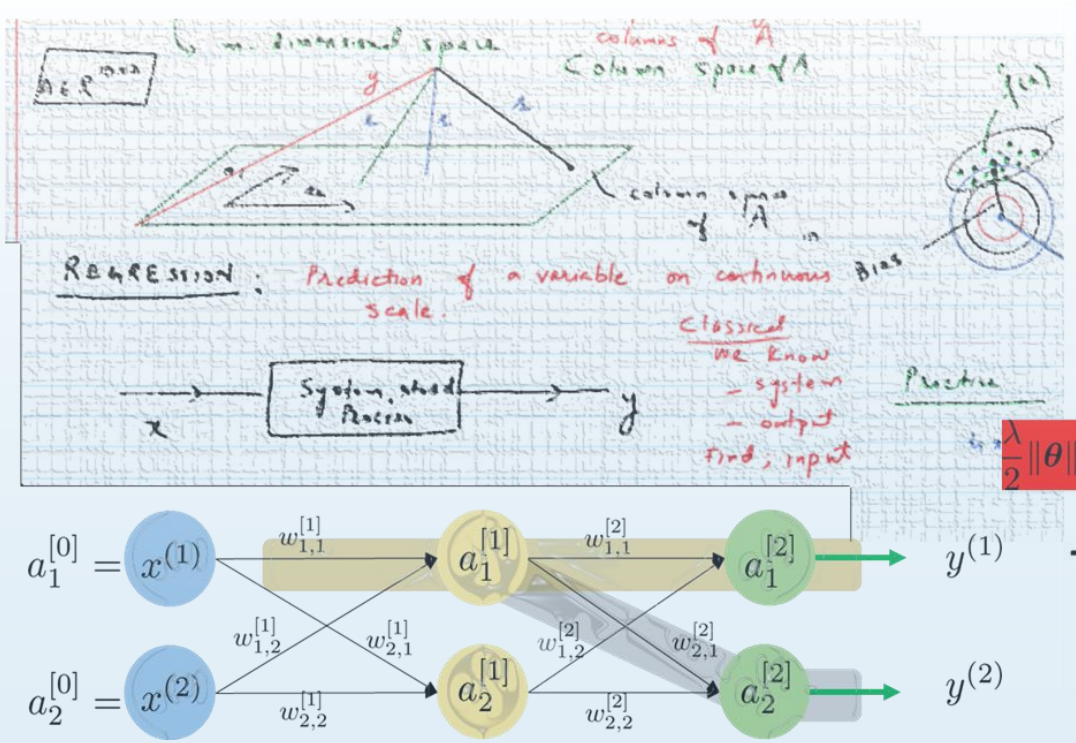
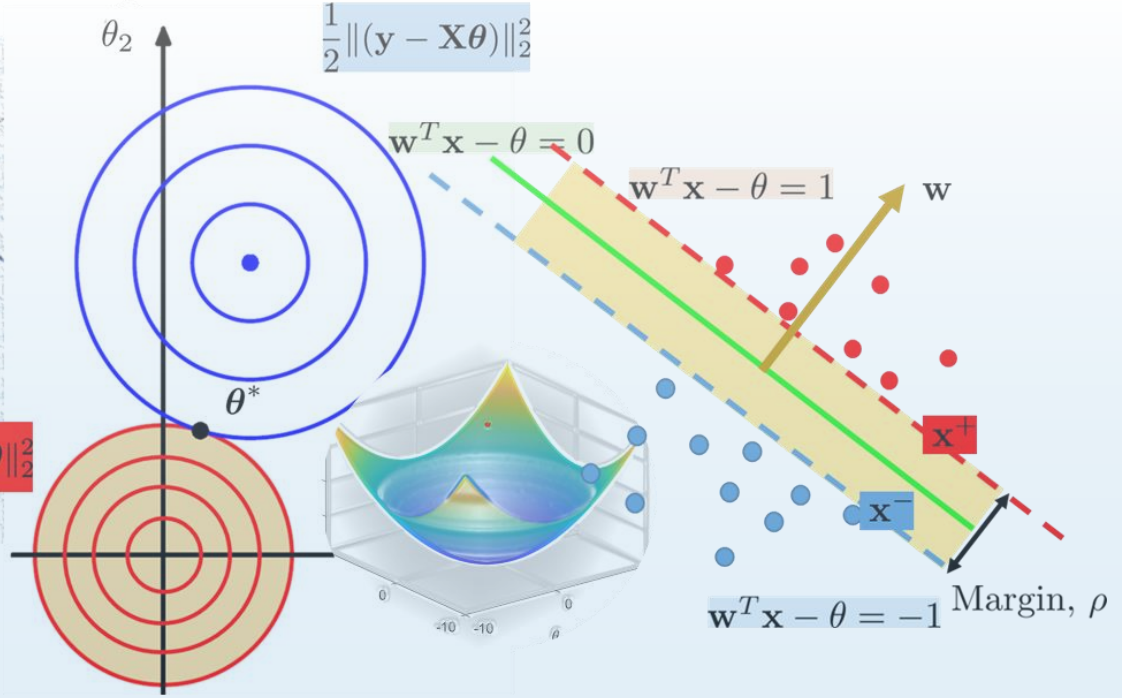




Machine Learning

Overview

Zubair Khalid

School of Science and Engineering

https://www.zubairkhalid.org/ee514_2025.html

LUMS

A Not-for-Profit University

About us!

Osama

Zubair

Hamza

Ahmad

What is this course about?

Introductory course in Machine Learning (ML) – Fundamental topics in

- Supervised learning
- Unsupervised learning
- Probabilistic Machine Learning

Course Objectives:

- To build mathematical foundations of ML and provide an appreciation for its applications
- To develop a comprehensive understanding of ML algorithms
- To provide an overview of probabilistic ML
- To provide experience in the implementation and evaluation of ML algorithms

Is this course a right choice for you?

Undergraduate students

- Interested in pursuing AI, Deep Learning and/or Machine Learning in their ***grad school***
- Interesting in pursuing a ***professional career*** focused on the development of Machine Learning solutions

Graduate students

- Want to do fundamental research in the area of AI
- Wish to apply AI in their research work

Modules

1- ML Overview

Course Overview, notation
Supervised Learning Setup

Weeks: 1

Components:

- Programming Assignment 0: Intro to Python, Setting up Environment

Modules

Classification
KNN
Evaluation Metrics, Curse of Dimensionality
Multi-class Classification

Weeks: 2,3

Components:

- Programming Assignment 1: KNN based
- Homework 1A

Modules

Linear Regression
Gradient Descent
Multi-variate Regression
Polynomial Regression
Bias-Variance Trade-off, Regularization

Weeks: 3,4

Components:

- Programming Assignment 2: Regression
- Homework 1B

Modules

4 – Bayesian Framework

Bayes Theorem
Naive Bayes Classification

Weeks: 5, 6

Components:

- Programming Assignment 3: Naïve Bayes Classifier
- Homework 2

Modules

Logistic Regression

Weeks: 6

Components:

- Programming Assignment 3: Logistic Regression
- Homework 2

Modules

6 – Perceptron and SVM

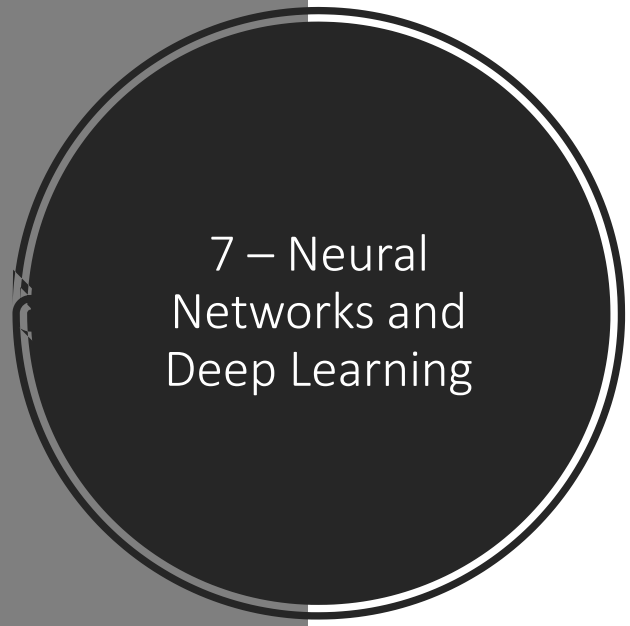
Perceptron Algorithm
SVM

Weeks: 7,8

Components:

- Programming Assignment 3: SVM
- Homework 2

Modules



Neural Networks (Feed-Forward Multilayer Neural Networks)
The Backpropagation Algorithm
Shallow vs Deep Architecture
Recurrent Neural Networks
Deep Generative Models

Weeks: 9,10,11

Components:

- Programming Assignment 4: Neural Networks
- Homework 3

Modules

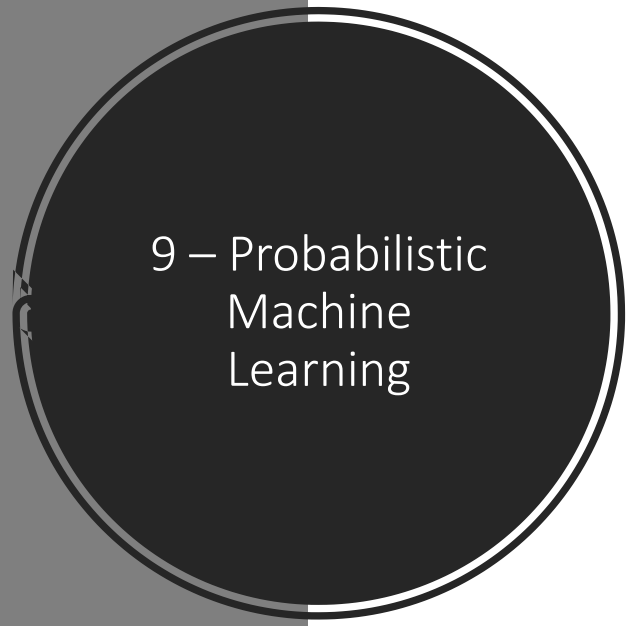
Unsupervised Learning Overview
Clustering (k-means)

Weeks: 12

Components:

- Homework 3

Modules



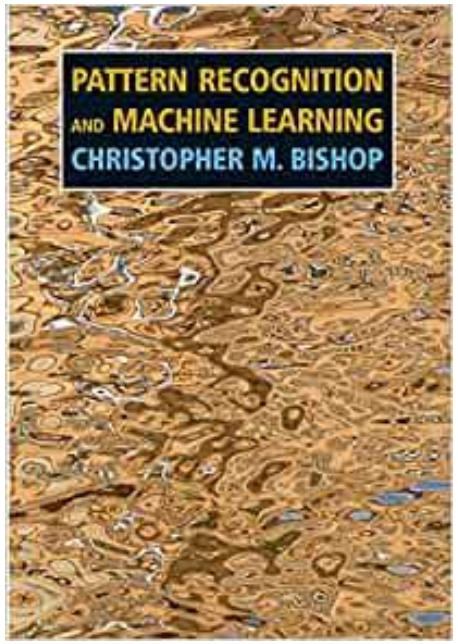
Probabilistic Machine Learning
Monte-Carlo Sampling Laplace Approximation MCMC methods Variational Auto-encoders

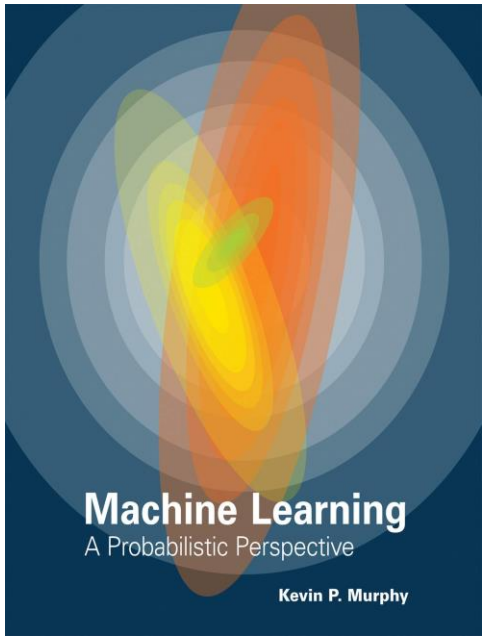
Weeks: 13,14

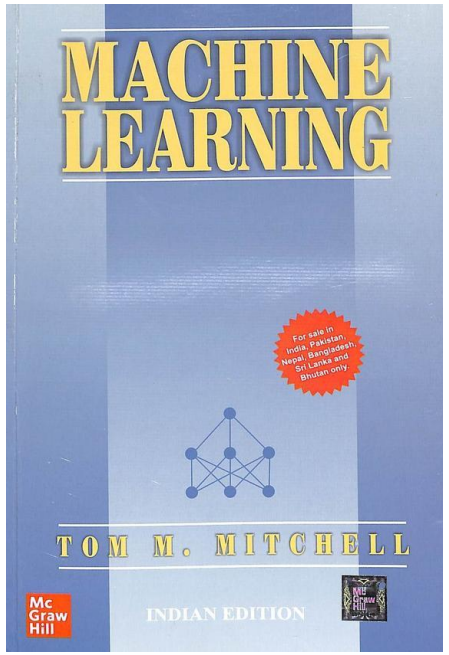
Components:

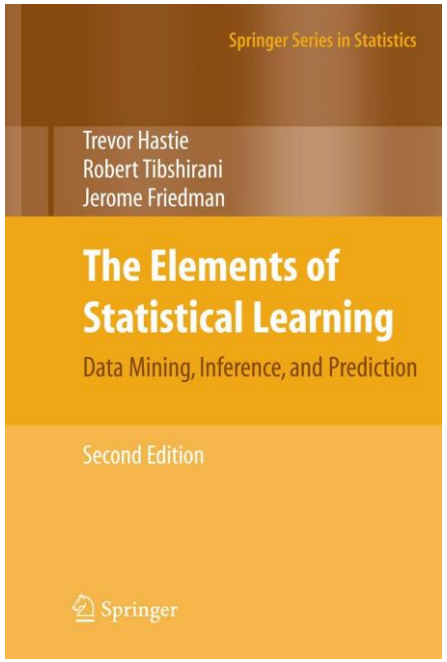
- Programming Assignment 5: Probabilistic Linear Regression and NNet
- Homework 3

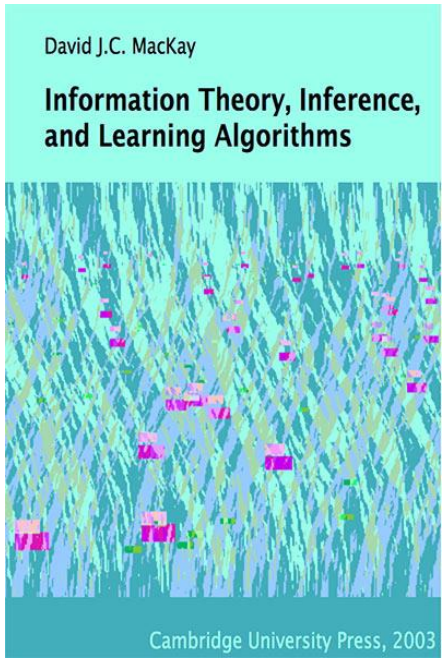
Suggested Reference Books











- (CB) Pattern Recognition and Machine Learning, Christopher M. Bishop
- (KM) Machine Learning: a Probabilistic Perspective, Kevin Murphy
- (TM) Machine Learning, Tom Mitchell
- (HTF) The Elements of Statistical Learning: Data mining, Inference, and Prediction, by Hastie, Tibshirani, Friedman
- (DM) Information Theory, Inference, and Learning Algorithms, David Mackay
- Lecture Notes/Slides will be shared.

Course Prerequisites

Undergraduate students

- Linear Algebra (MATH120)
- Probability (MATH230, DISC203, CS501)
- Programming (CS200, EE201)

Graduate students

- Encouraged to revise Linear Algebra and Probability concepts (on-the-fly)

We expect all the students to have good programming skills (in C/Python/MATLAB)

Note on Assignment 0!

Learning Interface

Communication:

Course Page: https://www.zubairkhalid.org/ee514_2025.html

Slack: Course-related questions or discussions. We will try to respond to the queries ASAP.

Office Hours: Posted on course page; distributed throughout the week

Email Policy:

Subject:

- 'ML-URGENT-Assignment Clarification'
- 'ML-NOT URGENT-Extend Assignment deadline'

Grading Distribution

- **Programming Assignments and Homeworks: 35%**
  - 5 Programming Assignments
  - 3 Homeworks
- **Quizzes: 15%** (Almost every week)
- **Project: 20%**
- **Final Exam: 30%**

Course Polices

- **Homework Late Policy**
  - 10% per day for 3 days. No submission after 3 days (72 hours)
- **Missed Quiz Policy**
  - No make-up for quiz
- **Plagiarism will be strictly dealt with as per university policies (take it seriously).**
- **Zero Tolerance for Plagiarism and Cheating**
- Re-grading can be requested after grade reporting, within the following time limits:
  - HW and Assignments: 2 days
  - Final Exam: 3 days

Course Polices

Harassment Policy

Harassment of any kind is **unacceptable**, whether it be sexual harassment, online harassment, bullying, coercion, stalking, verbal or physical abuse of any kind. Harassment is a very broad term; it includes both direct and indirect behaviour, it may be physical or psychological in nature, it may be perpetrated online or offline, on campus and off campus. It may be one offense, or it may comprise of several incidents which together amount to sexual harassment. It may include overt requests for sexual favours but can also constitute verbal or written communication of a loaded nature. Further details of what may constitute harassment may be found in the LUMS Sexual Harassment Policy, which is available as part of the university code of conduct.

LUMS has a Sexual Harassment Policy and a Sexual Harassment Inquiry Committee (SHIC). Any member of the LUMS community can file a formal or informal complaint with the SHIC. If you are unsure about the process of filing a complaint, wish to discuss your options or have any questions, concerns, or complaints, please write to the Office of Accessibility and Inclusion (OAI, oai@lums.edu.pk) and SHIC (shic@lums.edu.pk) —both of them exist to help and support you and they will do their best to assist you in whatever way they can.

To file a complaint, please write to harassment@lums.edu.pk.

Course Polices

Help related to equity and Belonging at SSE

SSE's Council on Equity and Belonging is committed to devising ways to provide a safe, inclusive, and respectful learning, living, and working environment for its students, faculty, and staff.
For help related to any such issue, please feel free to write to any member of the school council for help or feedback.

Mental Health Support at LUMS

-

For matters relating to counselling, kindly email student.counselling@lums.edu.pk, or visit https://osa.lums.edu.pk/content/student-counselling-office for more information.

You are welcome to write to me or speak to me if you find that your mental health is impacting your ability to participate in the course. However, should you choose not to do so, please contact the Counselling Unit and speak to a counsellor or speak to the OSA team and ask them to write to me so that any necessary accommodations can be made.

"As to methods, there may be a million and then some, but principles are few. The man who grasps principles can successfully select his own methods."

Ralph Waldo Emerson

Outline

- Course Overview, Logistics
- AI Overview
- Introduction to Machine Learning

AI Overview

What is Intelligence?

"Intelligence is the ability to **learn**, **understand**, and **apply** knowledge to adapt to new situations, reason, and solve problems."

- Human Intelligence
  - Problem-solving and reasoning
  - Learning from Experience
  - Adaptability
  - Emotional Intelligence

AI Overview

What is Artificial Intelligence (AI)?

"AI refers to the simulation of **human intelligence** in machines that are programmed to think and learn like humans."

- Narrow AI vs General AI
- SuperIntelligence

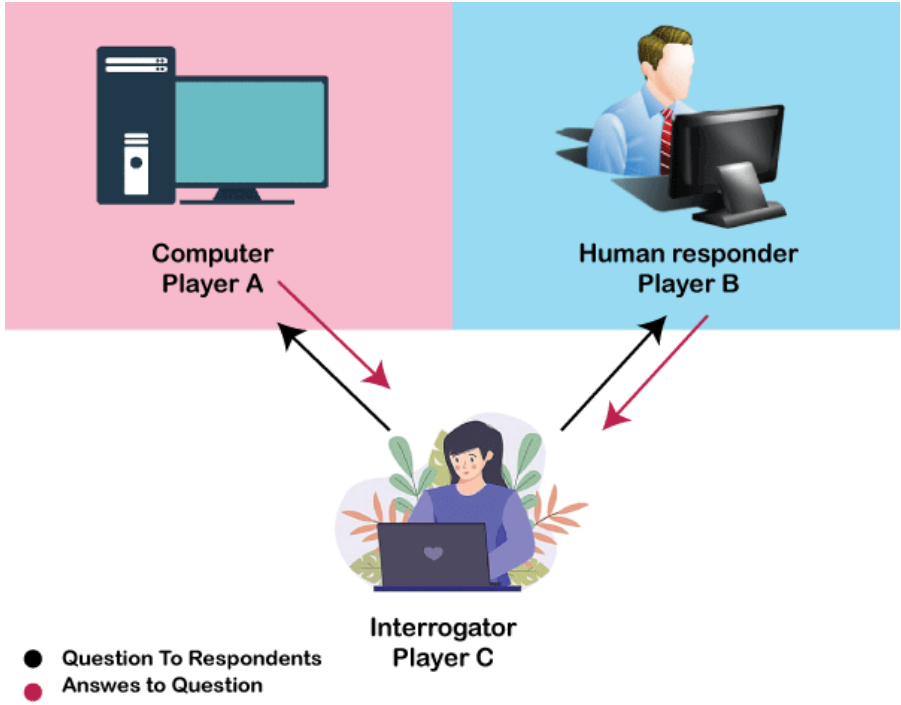
The term AI was coined in 1956 by John McCarthy at a conference.

AI Overview

History of AI – Turing Test

Alan Turing (1950) "Computing machinery and intelligence"

- **Can machines think or behave intelligently?**
- The Turing Test (imitation game) was designed as a way to judge the success or otherwise of an attempt to produce a thinking computer.
- Predicted that by 2000, a machine would have a 30% chance of fooling a person for 5 minutes
- What do you think?



AI Overview

History of AI – 1960s to 1990s

- Since the 1950s, much of the initial optimism surrounding AI has given way to a more realistic perspective.

- Key areas of focus include:
  - Machine learning
  - Multi-agent systems
  - Computer vision
  - Planning
  - Game Theory
  - Optimization, …

- 1997 (IBM's Deep Blue), 2012 (Deep Learning breakthrough).

Applications of AI

AI in Healthcare

- **AI-powered Diagnostics**

- AI systems can analyze medical images, such as X-rays, MRIs, and CT scans, to detect diseases like cancer, heart conditions, and more.

- These systems can identify patterns that might be missed by human eyes, resulting in earlier and more accurate diagnoses.

- **Example:** Google's DeepMind developed an AI system capable of diagnosing eye diseases with an accuracy that matches world-leading doctors.

Applications of AI

AI in Healthcare

- **Predictive Analytics**

- AI can analyze patient data and predict potential health risks, such as the likelihood of developing chronic diseases like diabetes or heart disease. This helps doctors create personalized treatment plans and intervene early.

- **Example:** IBM Watson Health analyzes vast datasets to predict disease progression and recommend personalized treatment options.

Applications of AI

AI in Healthcare

- **Robotic Surgery**

- AI-powered robotic systems assist in precision surgeries, allowing for less invasive procedures, faster recovery times, and reduced human error.

- **Example:** The da Vinci Surgical System uses AI to enhance the precision and control of surgeons during minimally invasive procedures.

Applications of AI

AI in Healthcare

- **Drug Discovery and Development**

- AI accelerates the drug discovery process by analyzing complex biological data to identify new drug candidates, significantly reducing the time and cost involved in developing new treatments.

- **Example:** AI was used by the pharmaceutical company Insilico Medicine to identify potential new compounds for treating fibrosis, which traditionally would have taken years.

Applications of AI

AI in Healthcare

- **Virtual Health Assistants**

- AI chatbots and virtual assistants provide personalized health advice, answer questions, and help patients manage chronic conditions by monitoring symptoms and offering medication reminders.

- **Example:** AI-driven platform allows patients to receive medical advice through their smartphones based on the symptoms they input.

Applications of AI

AI in Finance

- Fraud Detection
- Algorithmic Trading
- Risk Management
- Personalized Banking
- Credit Scoring and Loan Decisions
- Financial Forecasting

Applications of AI

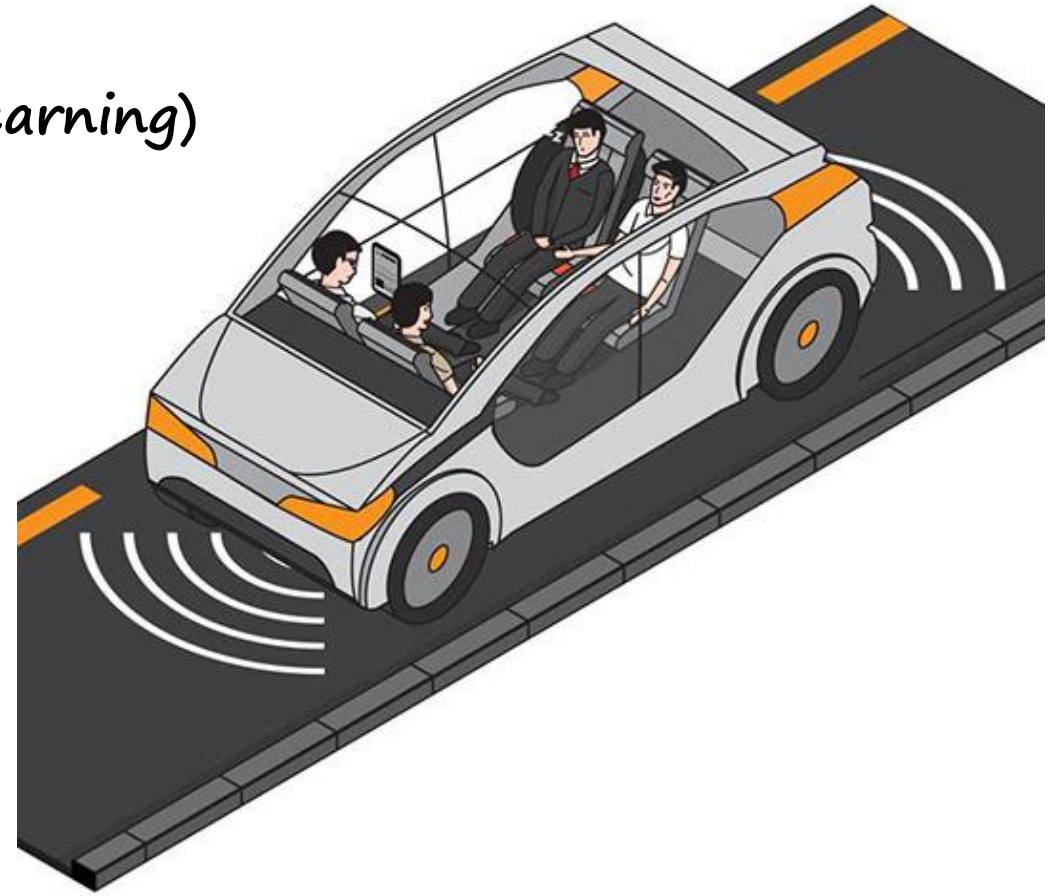
AI in Education

- Personalized Learning Systems
- Virtual Tutors and Chatbots
- AI in Assessments and Grading
- AI in Special Education

Applications of AI

AI in Autonomous Vehicles

- Deep Learning for Object Detection
- Decision Making and Planning (Reinforcement Learning)
- Simultaneous Localization and Mapping (SLAM)
- Path Planning and Navigation
- To name a few…



Applications of AI

AI in Retail

- Personalized product recommendations
- AI-powered chatbots and virtual assistants
- Inventory management and demand forecasting
- Dynamic pricing
- Visual search and image recognition
- Customer sentiment analysis

Applications of AI

AI in Agriculture

- Precision farming
- Crop monitoring and disease detection
- Soil health monitoring and analysis
- Automated machinery and robotics
- Predictive analytics and yield forecasting

Applications of AI

AI in NLP

- Virtual assistants (e.g., Siri, Alexa)
- Language translation
- Text generation and summarization
- Speech recognition and transcription
- Sentiment analysis
- Chatbots for customer support

Applications of AI

AI in Manufacturing

- Predictive maintenance
- Quality control and defect detection
- Supply chain optimization
- Autonomous robots and cobots (collaborative robots)
- Inventory management and optimization
- Production scheduling and process optimization

Applications of AI

AI in Urban Development

- Traffic management and optimization
- Smart energy grids and resource management
- Road safety and surveillance
- Waste management optimization
- Smart infrastructure and building management
- Environmental monitoring

Centre for Urban Informatics, Technology and Policy (CITY)

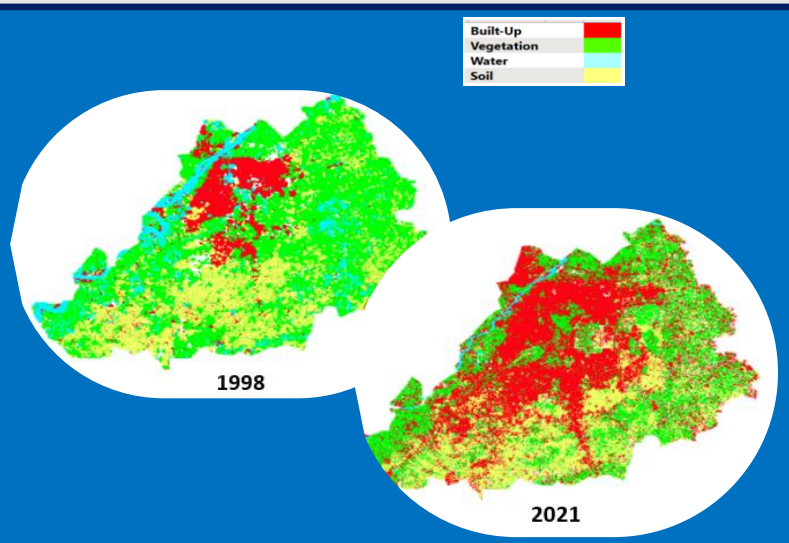
Saving our cities through technology and data-driven policy

VISION STATEMENT	*To become an internationally recognized regional hub of innovation in urban informatics, technology and data-driven policy-making for sustainable urban development.*

The Grand Challenge

Rapid urbanization, **resource scarcity** and **climate change** are testing the resilience of cities

Growth



Lahore's Built-up Area. 4.5X (1998 – 2022)

Mobility

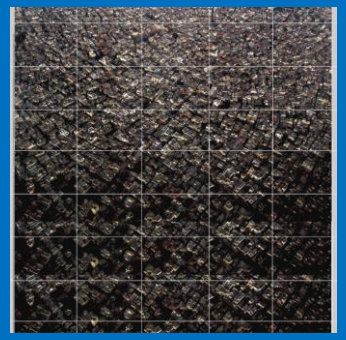
4x (2010-2024)

3x (2010-2024)

Environment and Health

CITY at LUMS

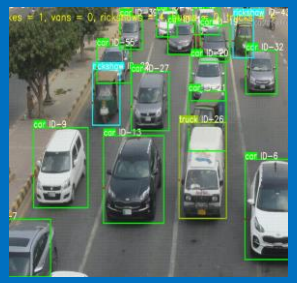
A center focusing on **climate-resilient urban solutions** through **technology, data, and policy integration**



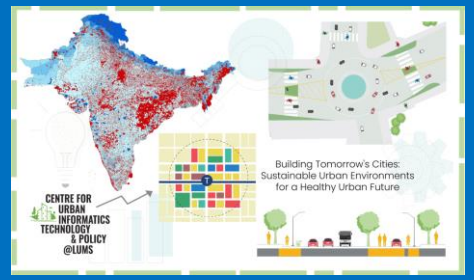
1 Research

Climate
Mobility
Space-time Resource Optimization

Solutions 2



Machine Vision & AI
Intelligent Traffic Insights
Data-Driven Policies



3 Outreach

Engagement with Government
International Collaborations, Grants
Seminars & Workshops

Academy 4

Workshops
Trainings
Urban Dictionary

Deep Tech
AI, Optimization, Modeling, Policy

Data Generation and Analysis

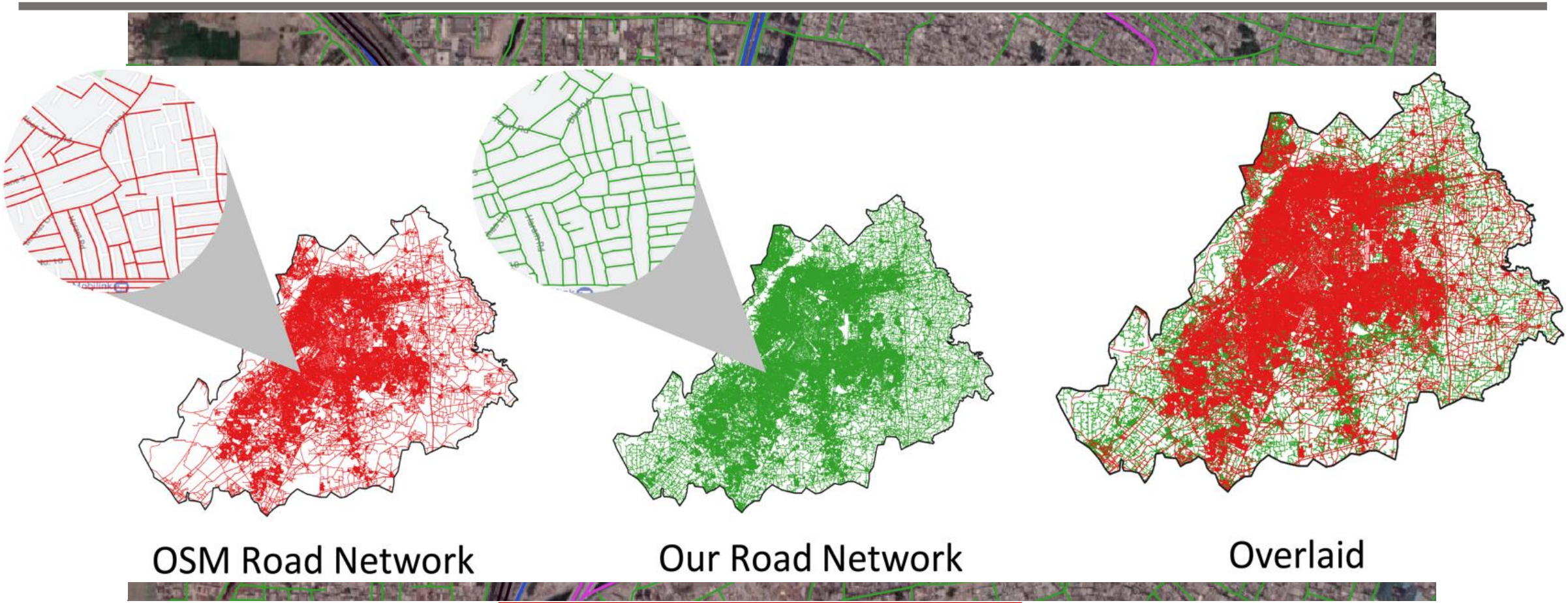
Evidence-Based Decision Making

Sustainable Communities

Inclusivity

Equity

CITY Research: Road Network



Analysis Tools:

QGIS, AI, and CV

Road Network

Center for Urban Informatics, Technology and Policy (CITY) at LUMS

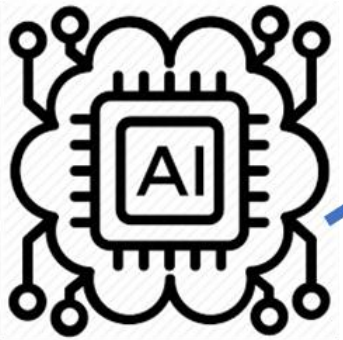
LUMS
A Not-for-Profit University

CITY Research: Mobility Informatics

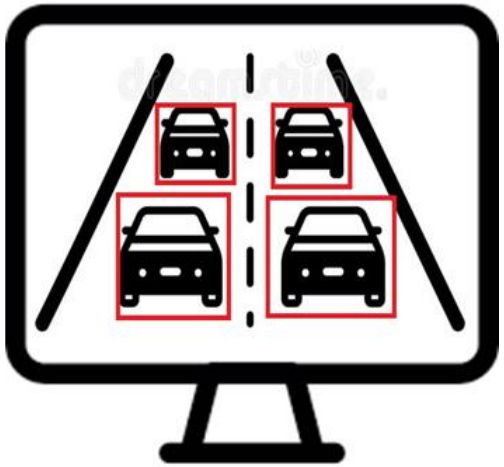
Road monitoring using camera

Database/Storage

AI model



Dashboard



AI based system for real time traffic flow density

Traffic Flows Density

CITY Research: Mobility Informatics

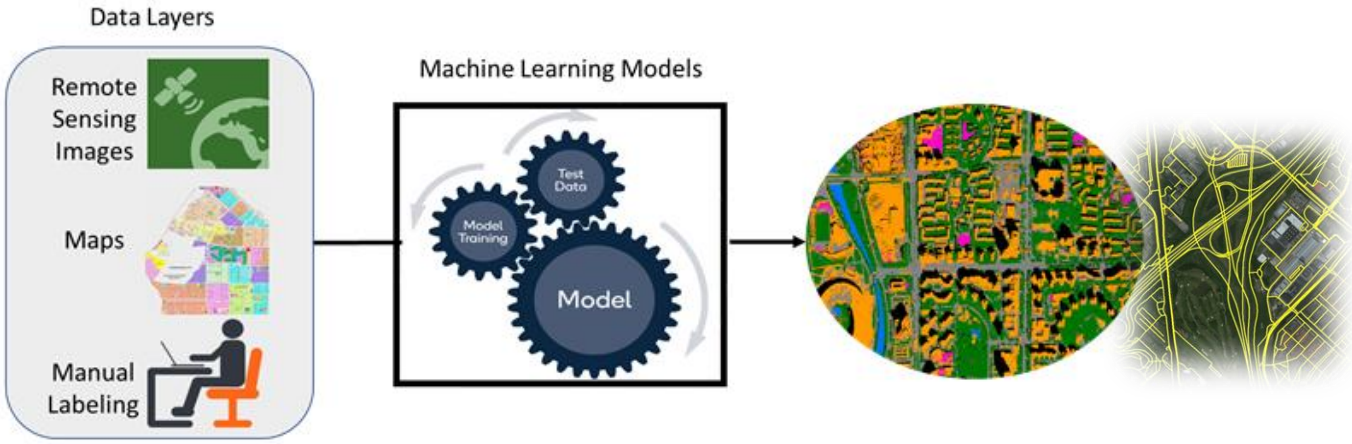
Prototype Sample

Traffic Flows Density

Center for Urb
Technology and Policy (CITY) at LUMS

LUMS
A Not-for-Profit University

CITY Research: Urban Sprawl



Land-Use Classification

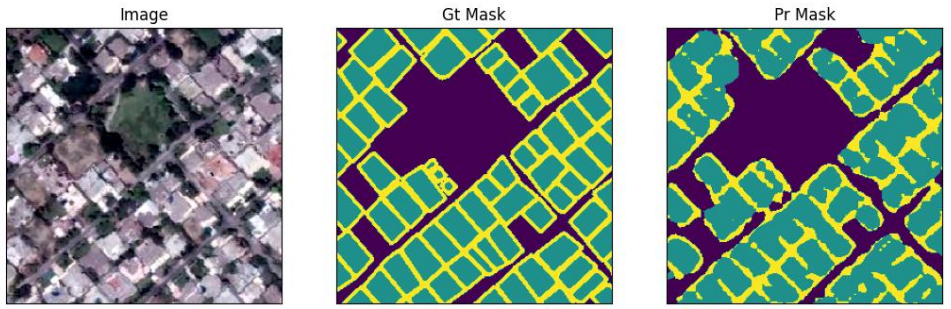
Building Footprint Detection

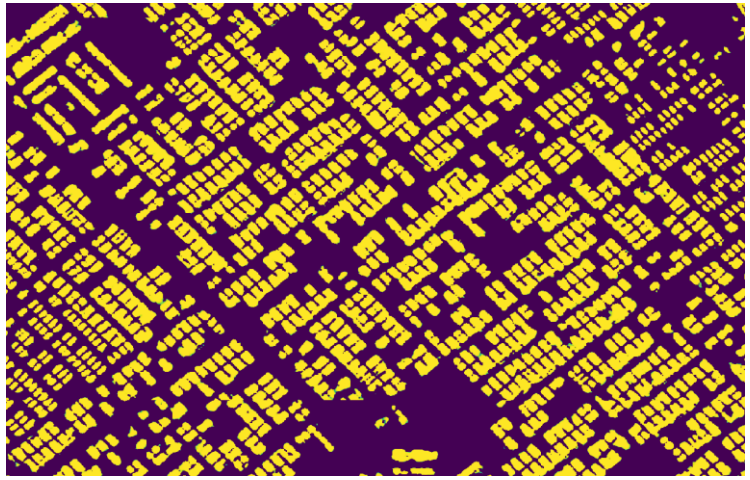
Transit-Oriented Development (Jain Mandir)

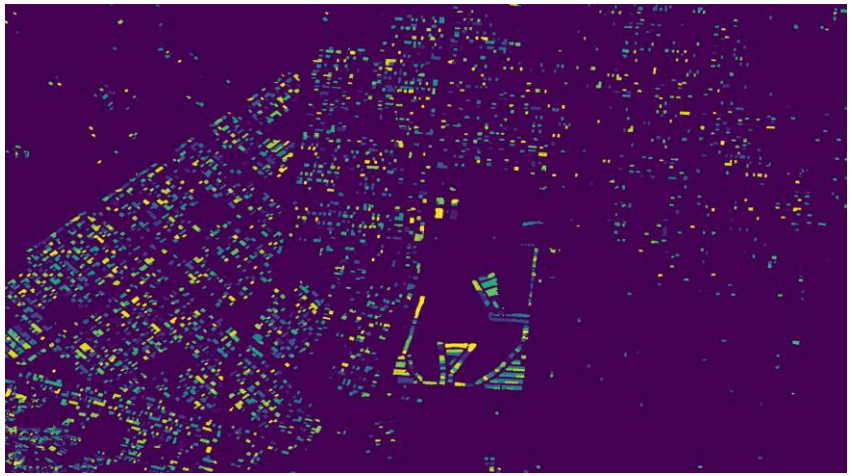
Spatiotemporal analysis of urban growth and land surface temperature

CITY Research: Urban Sprawl

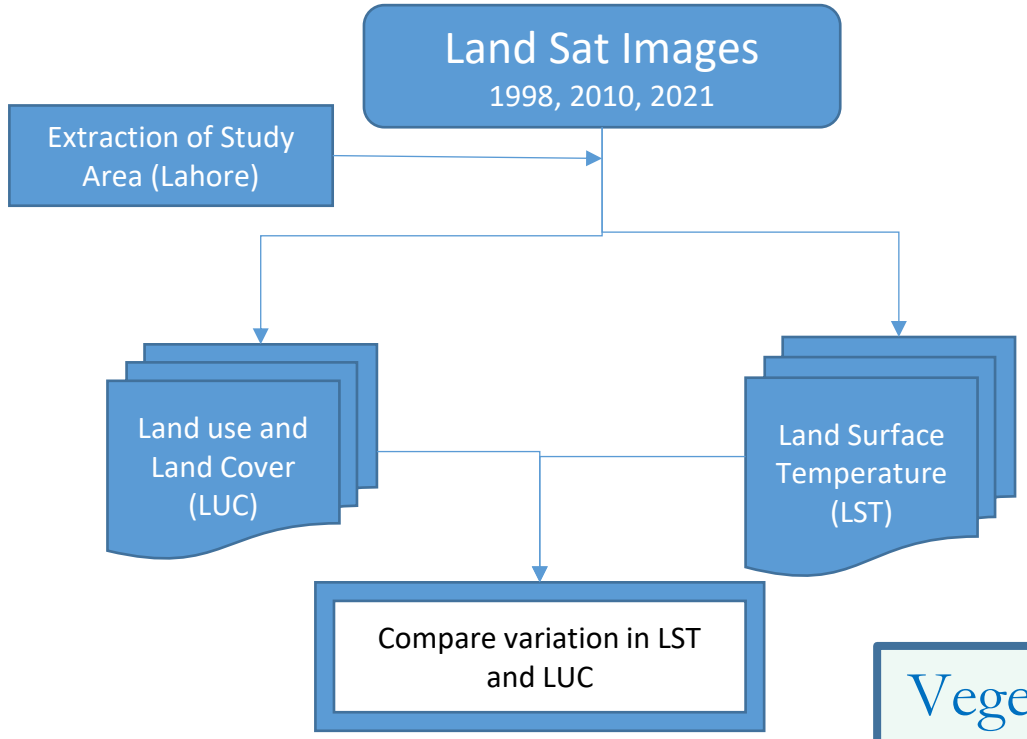
Automated Footprint Extraction

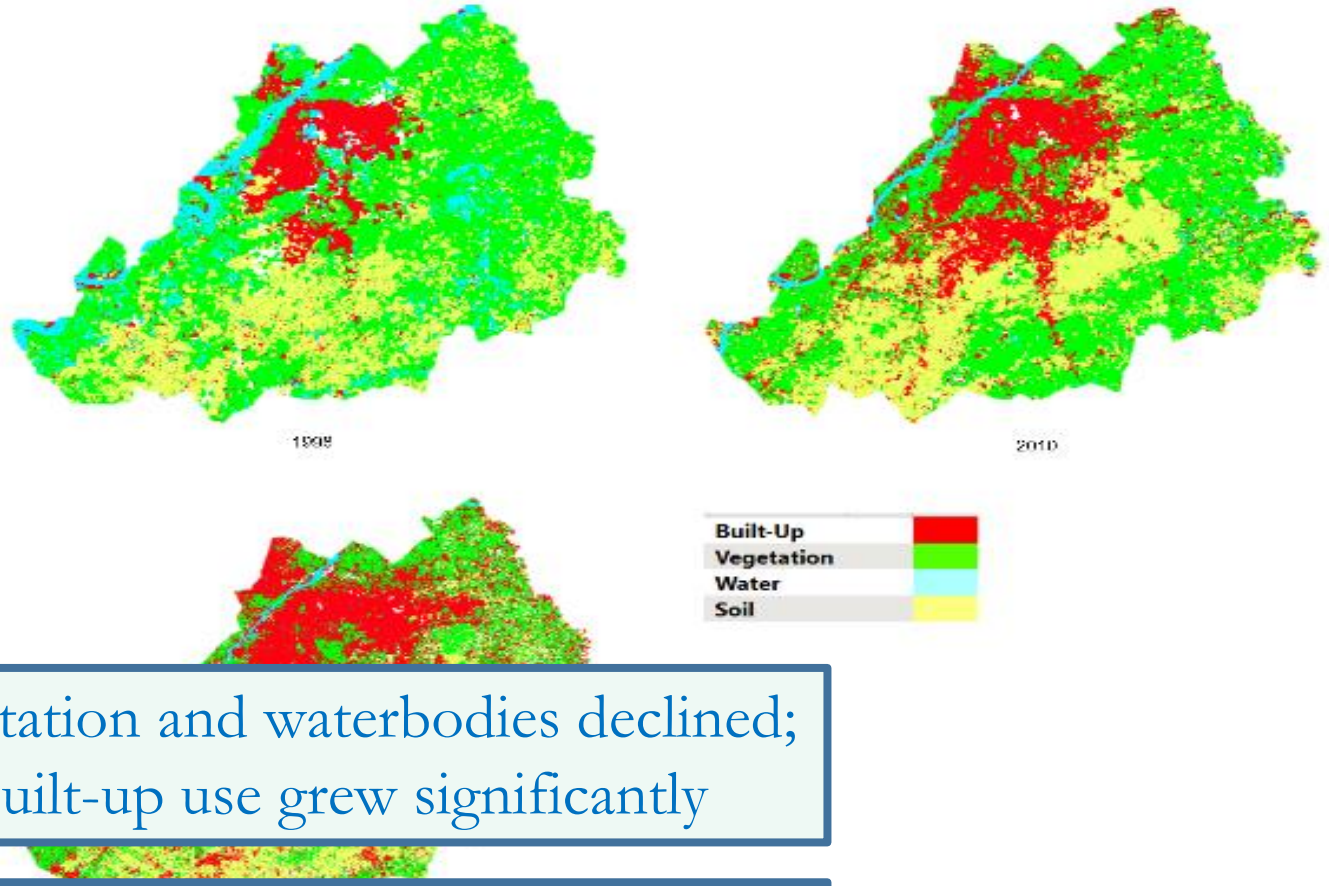






CITY Research: Urban Sprawl (LST)





Vegetation and waterbodies declined; Built-up use grew significantly

An average increase of 0.47 Celsius per year between 1998 and 2021

Center for Urban Informatics, Technology and Policy (CITY) at LUMS

LUMS
A Not-for-Profit University

CITY Solutions: Countering Air Pollution

Q: Where do we place 30 AQI Sensors?

- Historical AQI Data
- Points of Interests
- Meteorological Data
- Satellite Data (Aerosol)
- Road Network
- Population Density

↓

Dense Spatial Air Quality Inference

↓

Entropy Minimization Based Station Location Recommendation

↓

Location of New Station

CITY Solutions: Data-driven Urban Forestation

Problem: **Estimate green spaces and tree cover in urban areas**

Why?

Green spaces – significant contributors to ecosystem services: air purification and carbon sequestration

Challenges

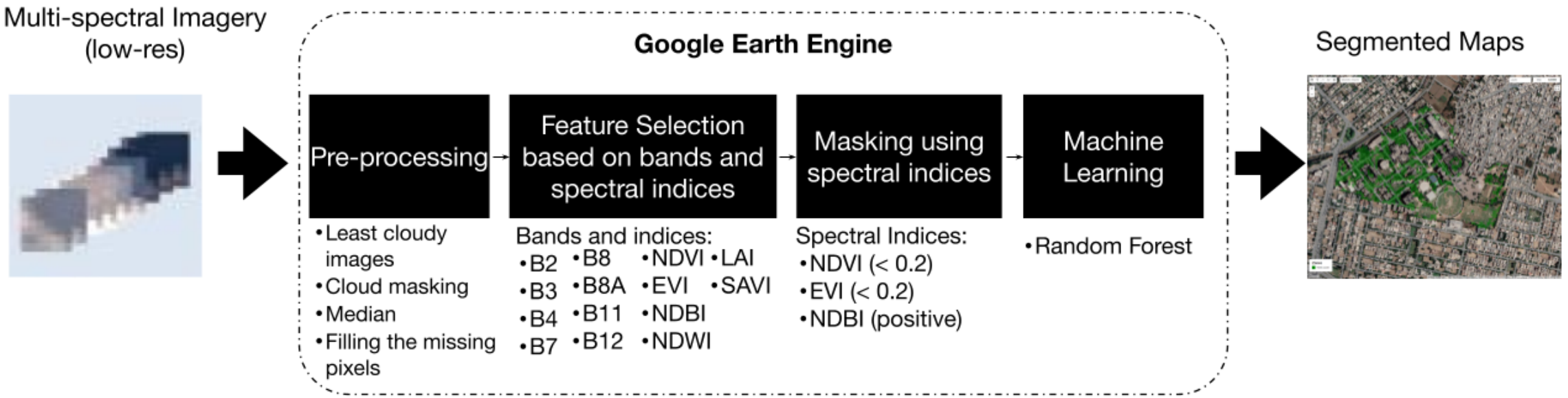
- Bottleneck: absence of spatially explicit data
- Huge variations in satellite imagery

Existing Methodologies

- Current methodologies focus on forest mapping
- A gap exists in urban tree cover estimation in developing countries with low resolution imagery

CITY Solutions: Data-driven Urban Forestation

Tree Cover Estimation using Machine Learning and Feature Selection on Sentinel-2 Data



CITY Solutions: Data-driven Urban Forestation

LUMS Ground Truth

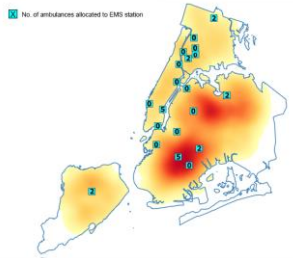
Proposed

ESA WorldCover

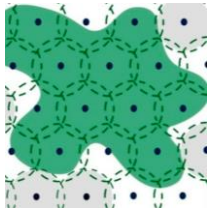
Model	Pred. area (acres)	Masking	Spectral indices	Pixel-wise Test Accuracy (%)	Kappa Score
RF-spectral-bands	29.5	No	No	0.93	0.81
RF-spectral-indices	28	No	Yes	0.95	0.88
Proposed	25	Yes	Yes	**0.99**	**0.92**
ESA WorldCover Product [5]	16	-	-	0.74	-
DeepLabv3 [6]	28	No	No	0.80	-

CITY Solutions: Improving Emergency Response Time

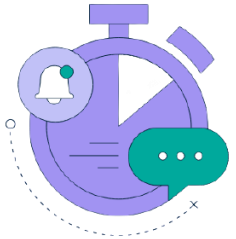
Improving health facilities and safety of inhabitants of urban areas of Pakistan using data-driven methods



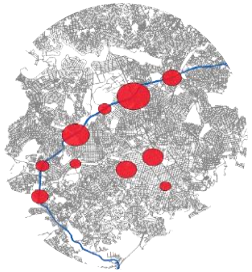
Optimized Allocation of EMS Resources



Maximizing Coverage



Minimize Response Time



Accidents Hot-spot Analysis

Real-time Route Planning

Center for Urban Informatics, Technology and Policy (CITY) at LUMS

LUMS
A Not-for-Profit University

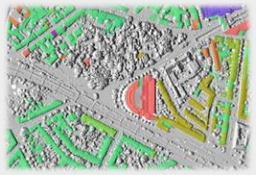
Urban Health: Emergency Response Services

Data Layers

Rescue 1122 Data

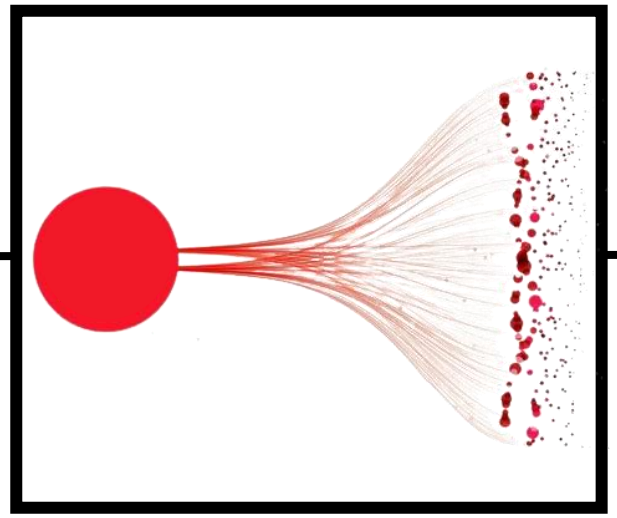
RESCUE 1122

GIS Data

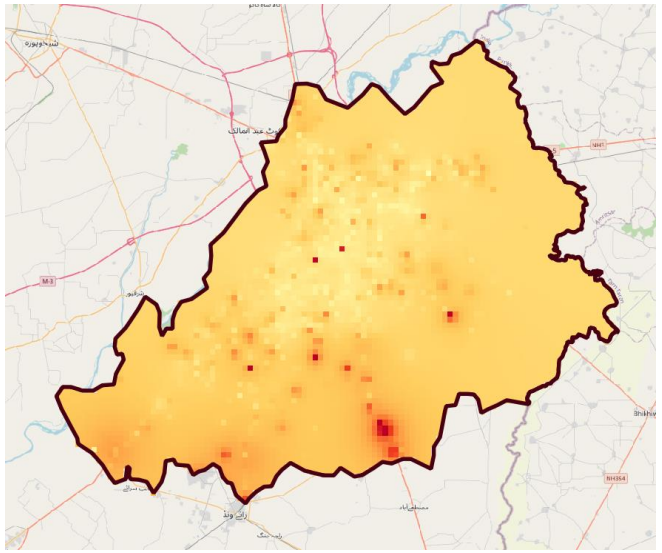


Google Data

Analytical Modeling and Data Analytics



Information Insights

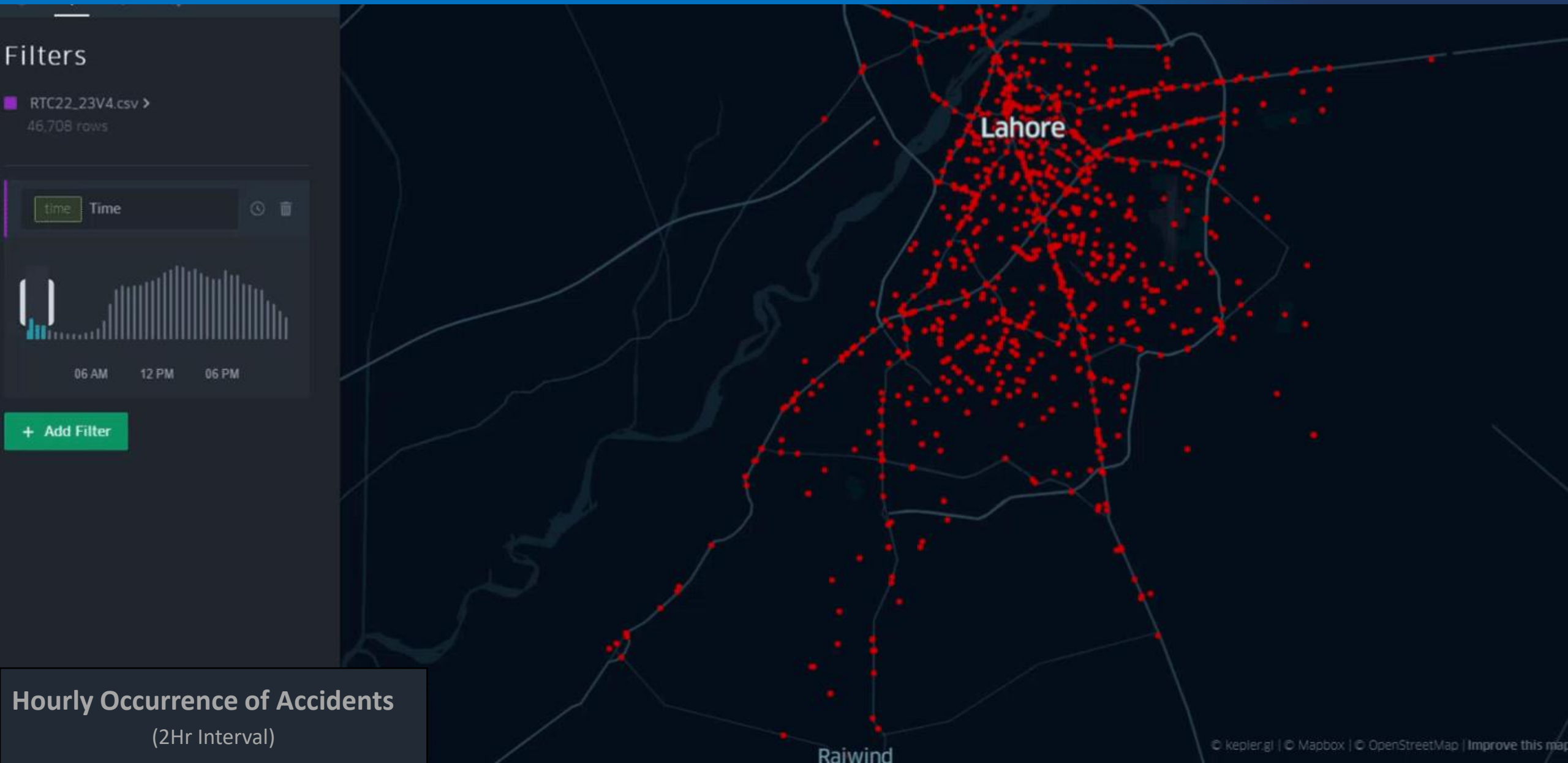


Prototype Study:
Traffic accident heatmaps and ambulance allocation for different scenarios for Lahore

Center for Urban Informatics, Technology and Policy (CITY) at LUMS

LUMS
A Not-for-Profit University

Spatio-temporal Distribution of Accidents



Hourly Occurrence of Accidents

(2Hr Interval)

Spatio-temporal Distribution of Accidents

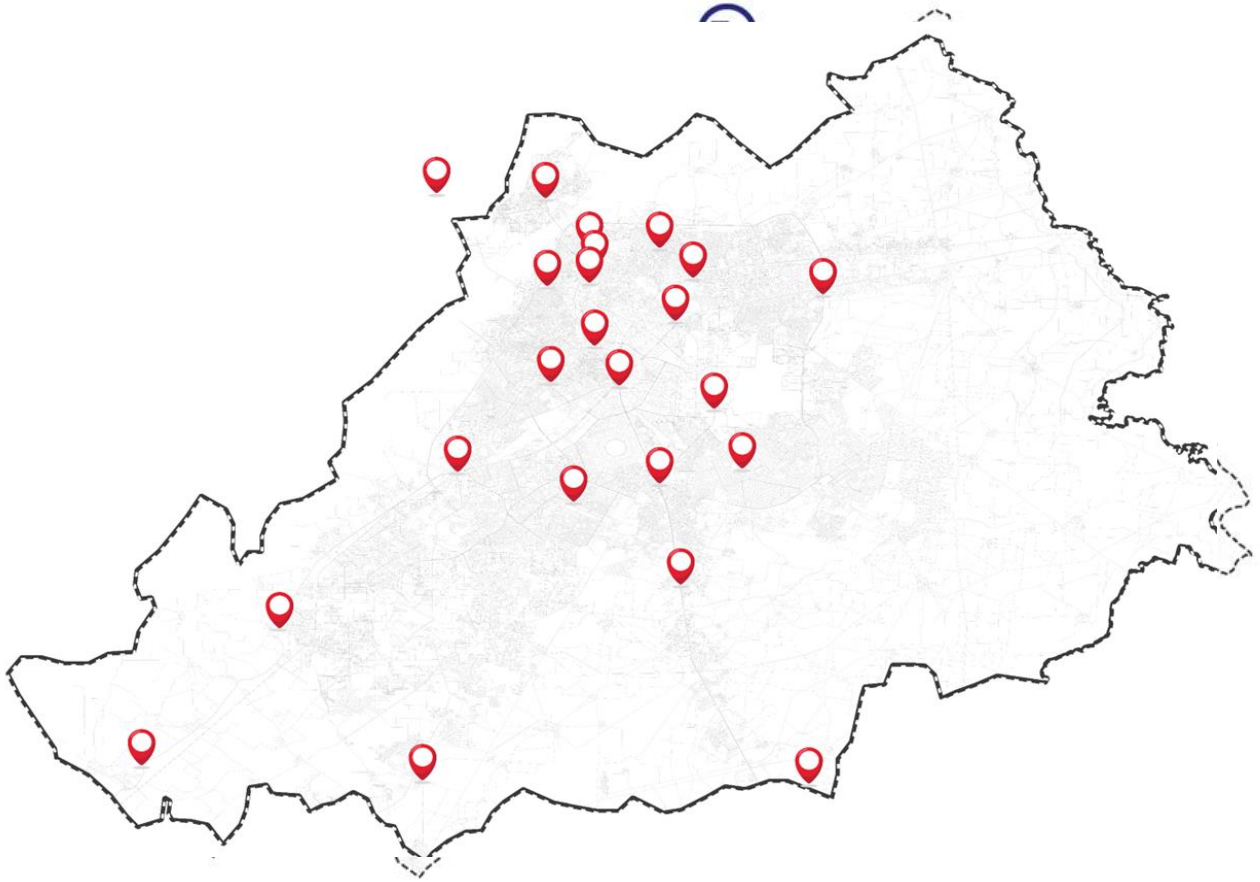
Category	Accidents	% Share	Total Accidents
Bike	28651	61.34	46709
Shifted/Dead	2658	5.69	
Pedestrian	15150	32.43	

High Se
Motorcy
Pedestr

61.34 % Accidents Involved Bikes

32.43% Accidents Involved Pedestrians

5.69% Accidents Resulted in Critical Injury



RTC Blackspots

Center for Urban Informatics, Technology and Policy (CITY) at LUMS

LUMS
A Not-for-Profit University

Outline

- Course Overview, Logistics
- AI Overview
- Introduction to Machine Learning

Machine Learning Overview

What is Machine Learning?

"The activity or process of gaining knowledge or skill by studying, practicing, being taught, or experiencing something."

Merriam Webster dictionary

*"A computer program is said to learn from experience **E** with respect to some class of tasks **T** and performance measure **P**, if its performance at tasks in **T**, as measured by **P**, improves with experience **E**."*

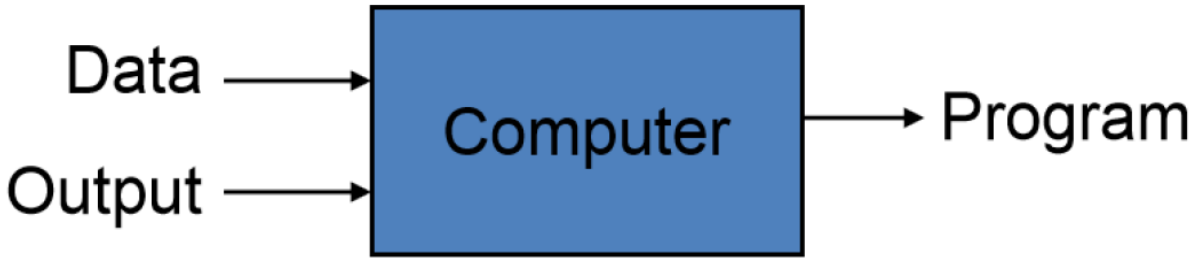
Tom Mitchell

Machine Learning Overview

What is Machine Learning?

- Automating the process of automation
- Getting computers to program themselves

Traditional Programming

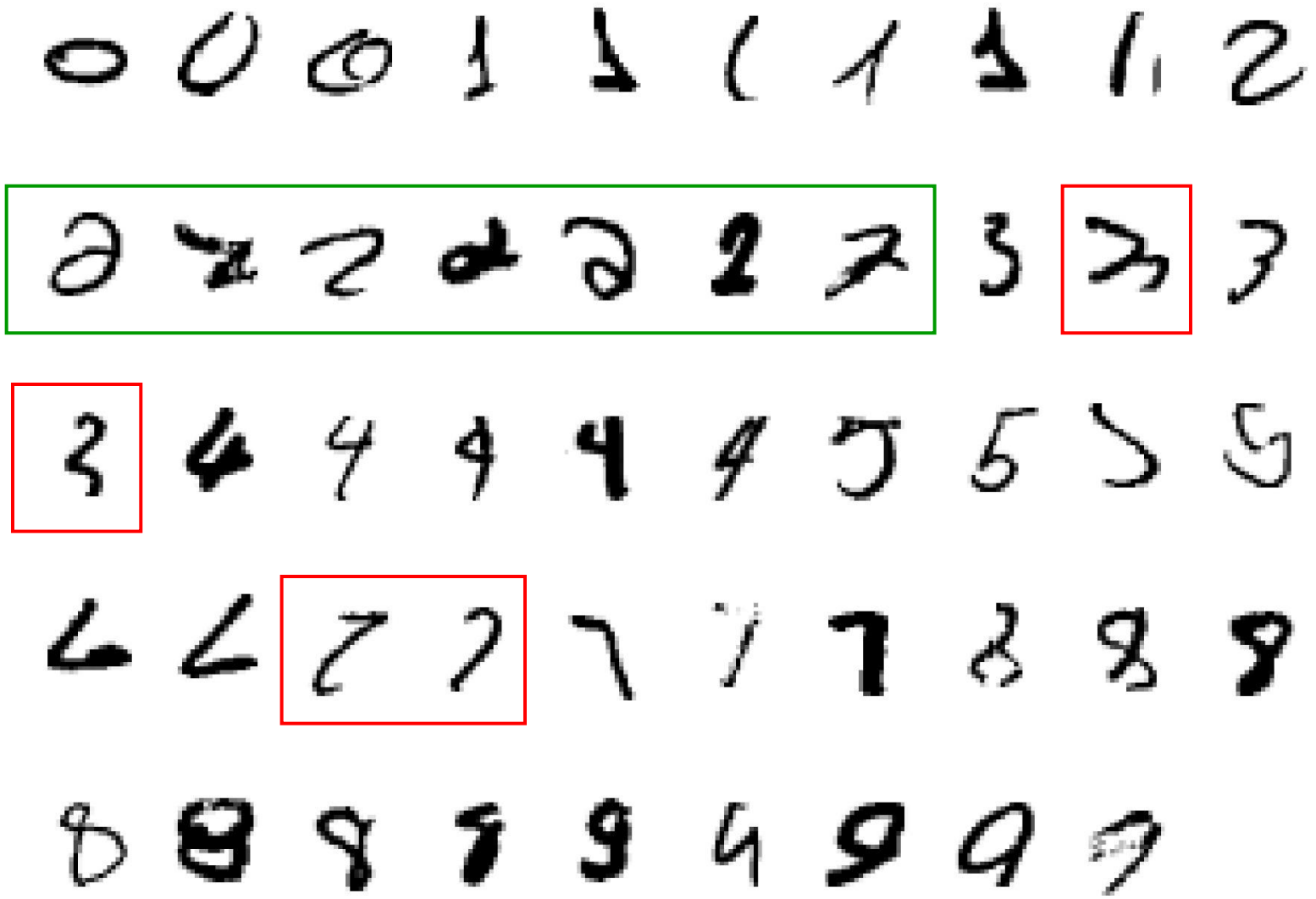


Machine Learning

Given examples (training data), make a machine learn system behavior or discover patterns

Machine Learning Overview

Classical Example: Recognize hand-written 2!

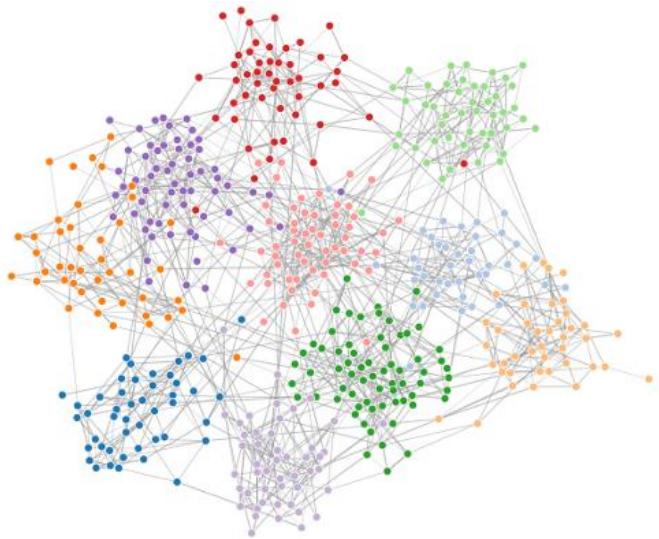


Machine Learning: Overview

What is Machine Learning?

Given examples (training data), make a machine learn system behavior or discover patterns

Data



Algorithm

Model

$$f(\mathbf{x})$$

Given to us

We need to design it

Final output which enables us to make predictions

Machine Learning: Overview

Algorithms vs Model

- Linear regression algorithm produces a model, that is, a vector of values of the coefficients of the model.

- Decision tree algorithm produces a model comprised of a tree of if-then statements with specific values.

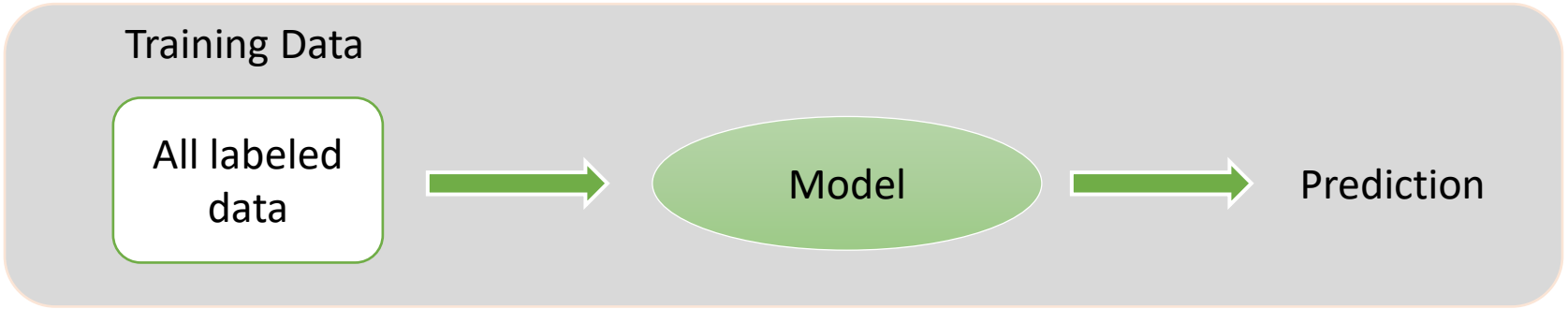
- Neural network along with backpropagation + gradient descent: produces a model comprised of a trained (weights assigned) neural network.

Machine Learning: Overview

Nature of ML Problems

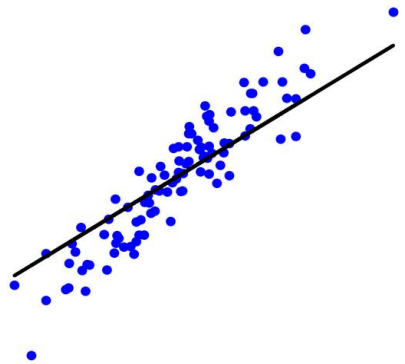
1. Supervised Learning

The learning algorithm would receive a set of inputs along with the corresponding correct outputs to train a model



Supervised Learning

Regression



Regression: Quantitative Prediction on a continuous scale

Examples: Prediction of

- Age of a person from his/her photo
- Price of 10 Marla, 5-bedroom house in 2050
- USD/PKR exchange rate after one year
- Efficacy of vaccine or medicine
- Average temperature/Rainfall during monsoon
- Cumulative score in EE514/CS35 course
- Probability of decrease in the electricity prices in Pakistan
- No. of steps per day

What do all these problems have in common?

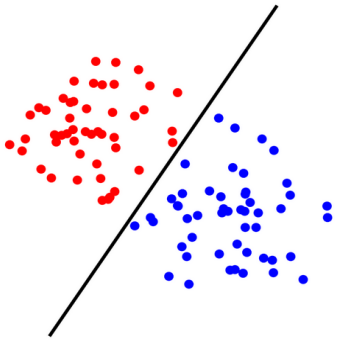
Continuous outputs

Predicting continuous outputs is called regression

Supervised Learning

Classification

Classification: Given a data sample, predict its class **(discrete)**



Examples: Prediction of

- Gender of a person using his/her photo or hand-writing style
- Spam filtering
- Object or face detection in a photo
- Temperature/Rainfall normal or abnormal during monsoon
- Letter grade in EE514/CS535 course
- Decrease expected in electricity prices in Pakistan next year
- More than 10000 Steps taken today

What do all these problems have in common?

Discrete outputs: Categorical

Yes/No (Binary Classification)

Multi-class classification: multiple classes

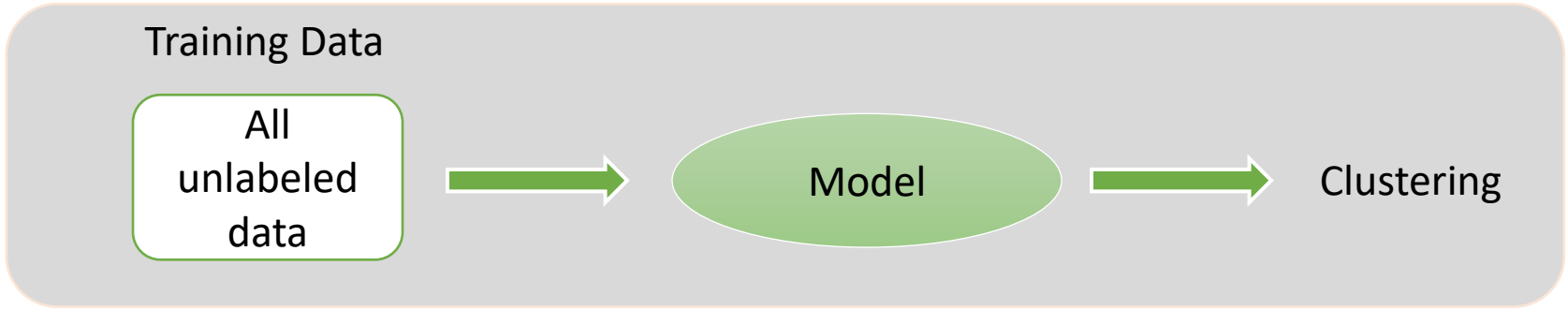
Predicting a categorical output is called classification

Machine Learning: Overview

Nature of ML Problems

2. Unsupervised Learning

The learning algorithm would receive unlabeled raw data to train a model and to find patterns in the data

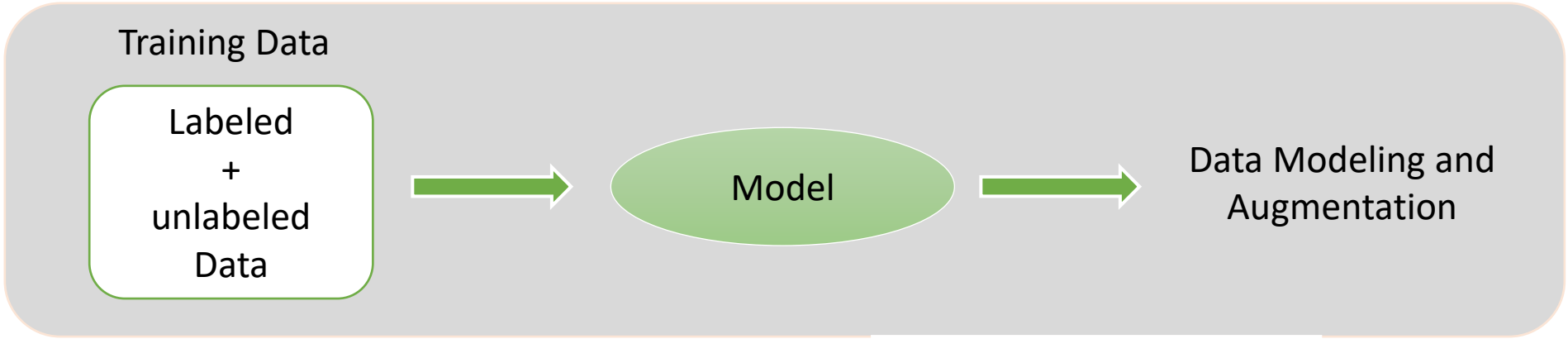


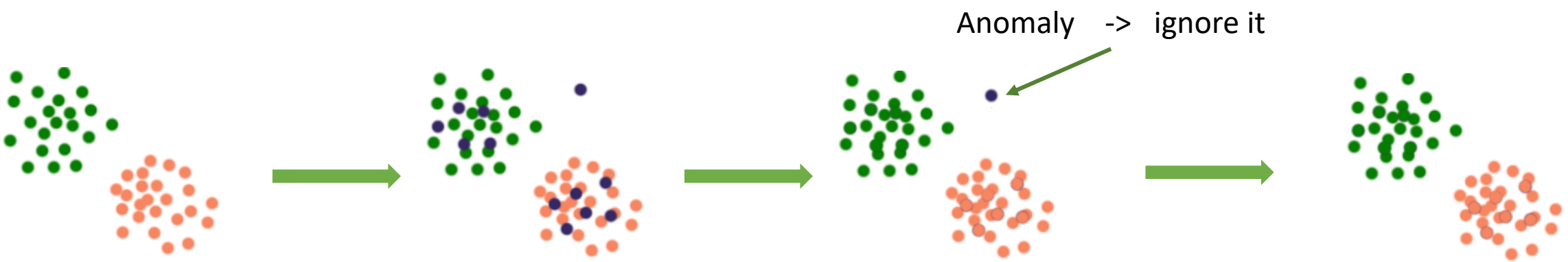
Machine Learning: Overview

Nature of ML Problems

3. Semi-supervised Learning

- *The learning algorithm receives labeled and unlabeled raw data to train a model*
- *Main objective is to efficiently accommodate the unlabeled data*





Use labeled data to build a model

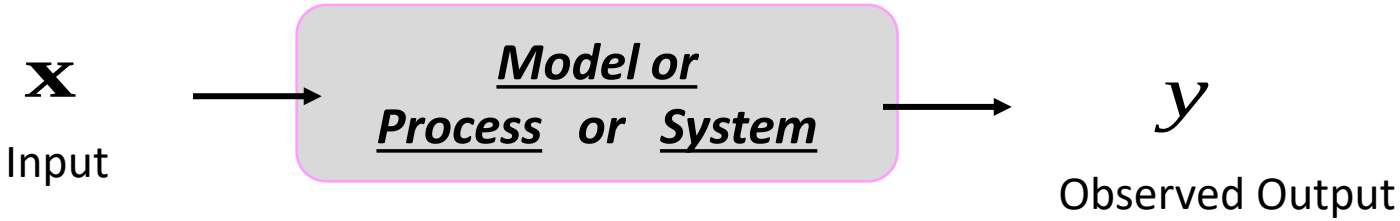
Place unlabeled data with model

Use the model to label the unlabeled data

Fit the model again with the combined data

Machine Learning: Overview

Training Data Collection

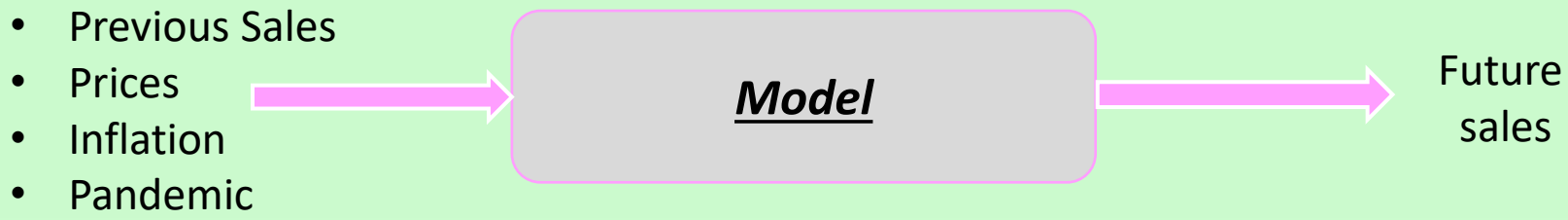


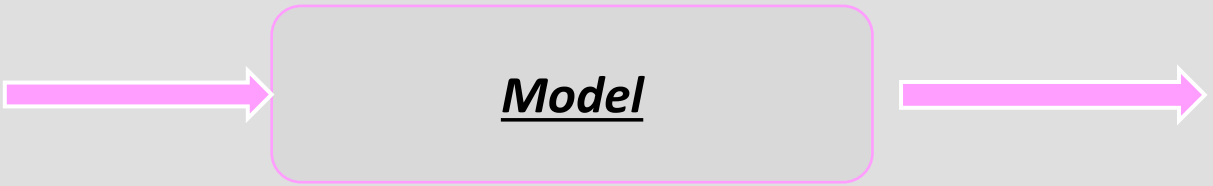
PROCESS or ***SYSTEM*** : Underlying physical or logical phenomenon which maps our input data to our observed output

Collect the training data by observing our unknown **PROCESS** or **SYSTEM**

Machine Learning: Overview

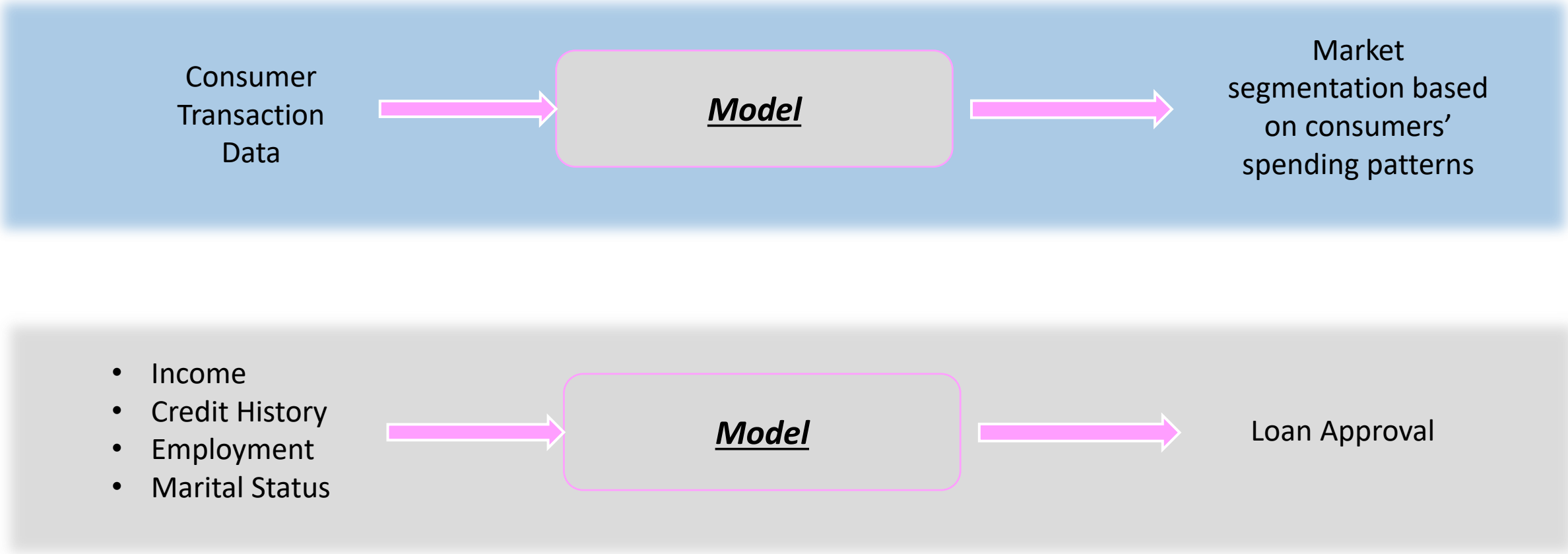
Example Systems





Machine Learning: Overview

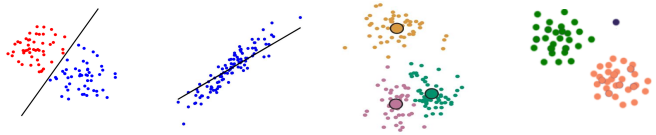
Example Systems



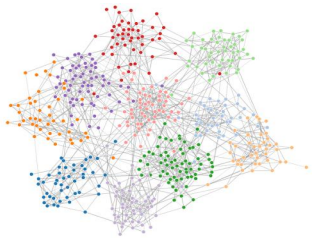
Machine Learning: Overview

Typical Flow

Problem Nature Identification



Training Data Collection



Model Training

$$f(\mathbf{x})$$

Model Validation

Model Deployment

Supervised Learning Setup

Reference:

- CB: sec 1.1
- HTF section 2.1
- KM: sec. 1.1, 1.2