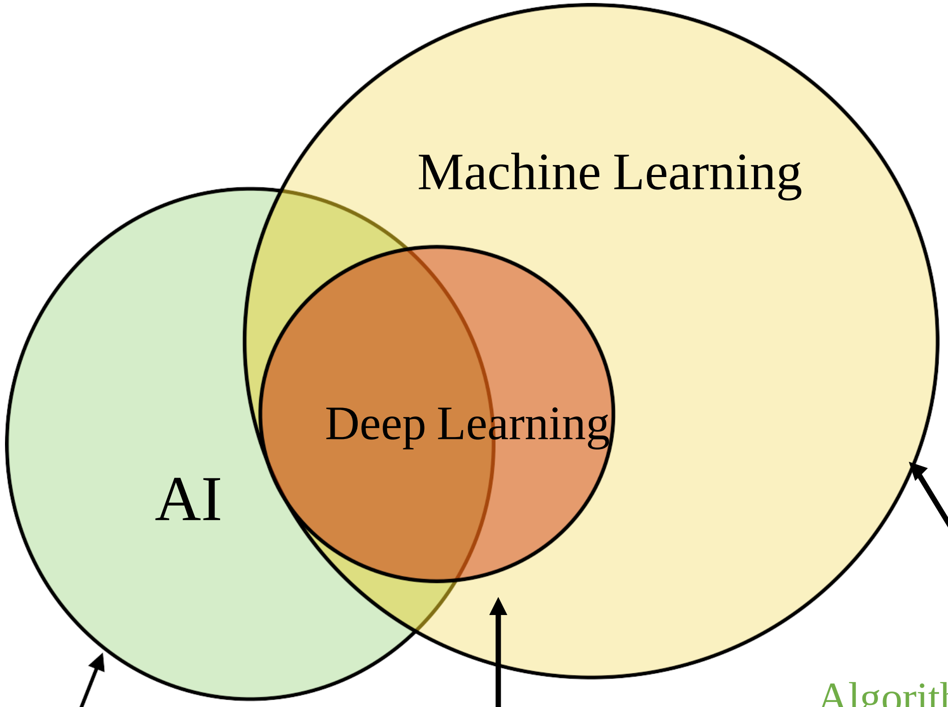
INDIAN CITIES OF THE FUTURE

Tanujit Chakraborty
Sorbonne University
POMS - XLRI

AI-ML and their impact in Geoscience



A non-biological system that possesses intelligence through rules

Algorithms that parameterize multilayer neural networks and then learn representations of data with multiple layers of abstraction

Algorithms that learn models/representations/ rules automatically from data/examples

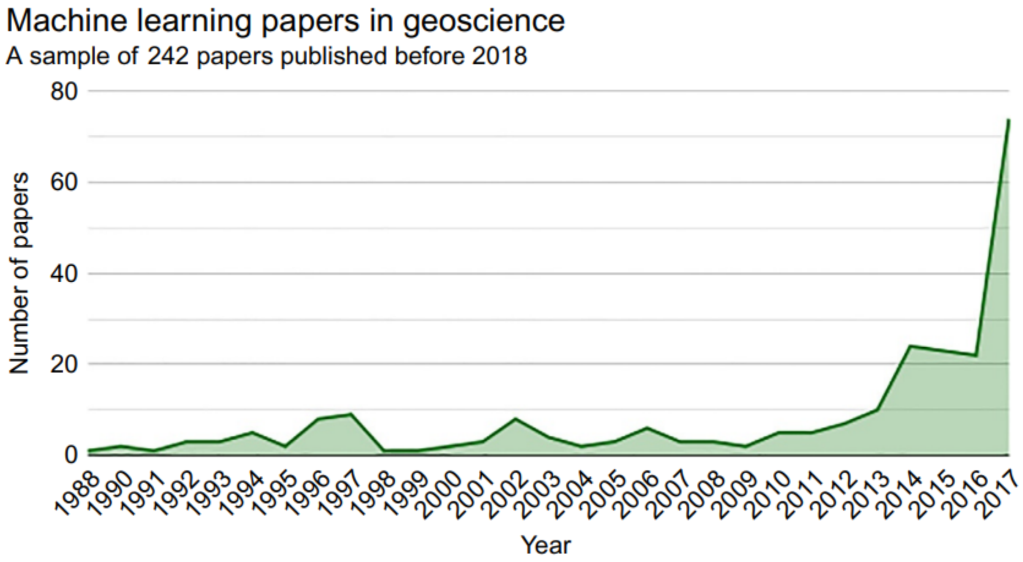


Fig. Bibliometry of 242 papers in machine learning for geoscience per year. Search terms include variations of machine learning terms and geoscientific subdisciplines but exclude remote sensing and kriging.

Dramsch, J. S. (2020). 70 years of machine learning in geoscience in review. Advances in geophysics, 61, 1-55.

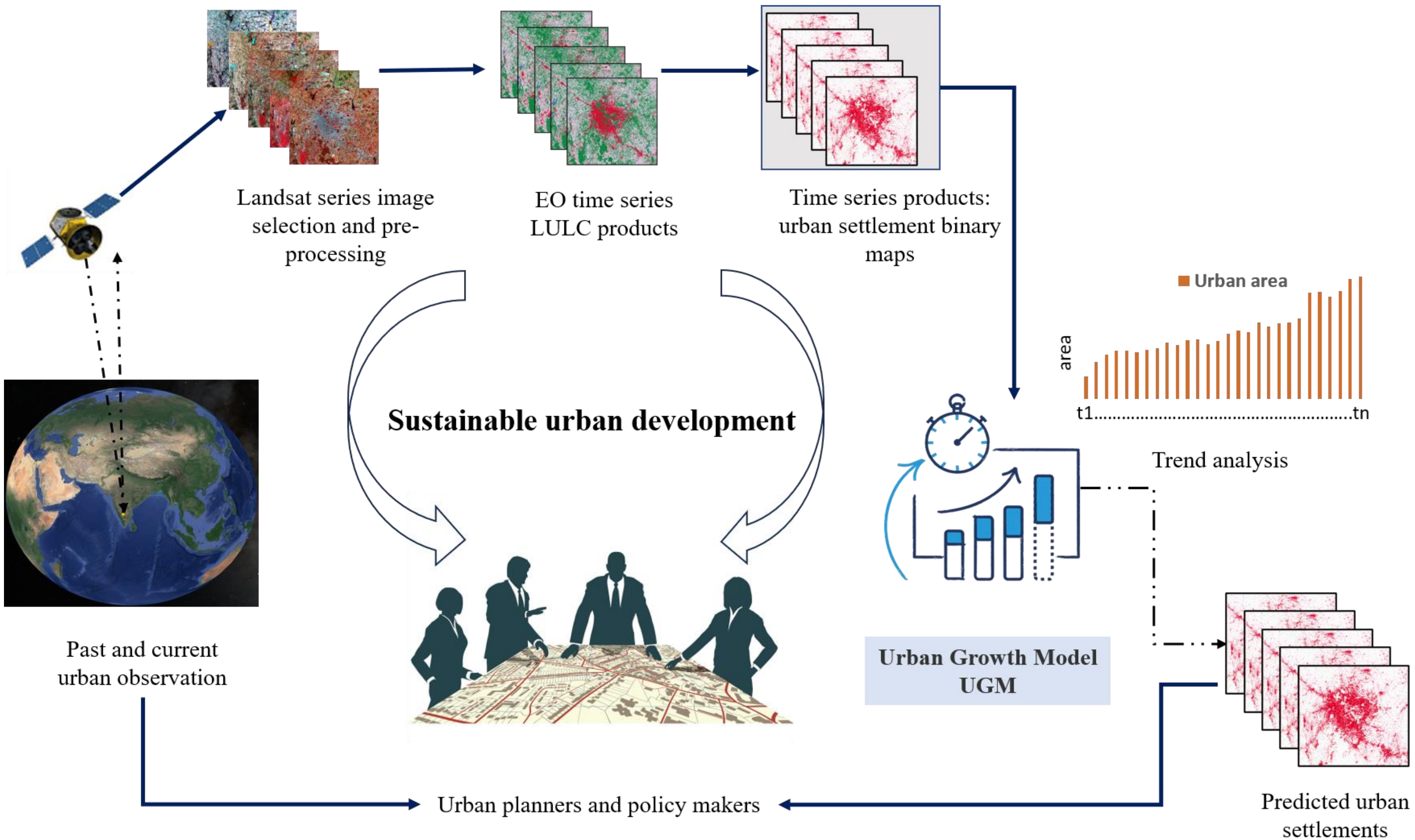
AI meets Urban Planning

- Urban planning is a multi-faceted discipline in which planners collect data points from targeted areas to determine if a store, hospital, road networks or other types of buildings are likely needed in that area.

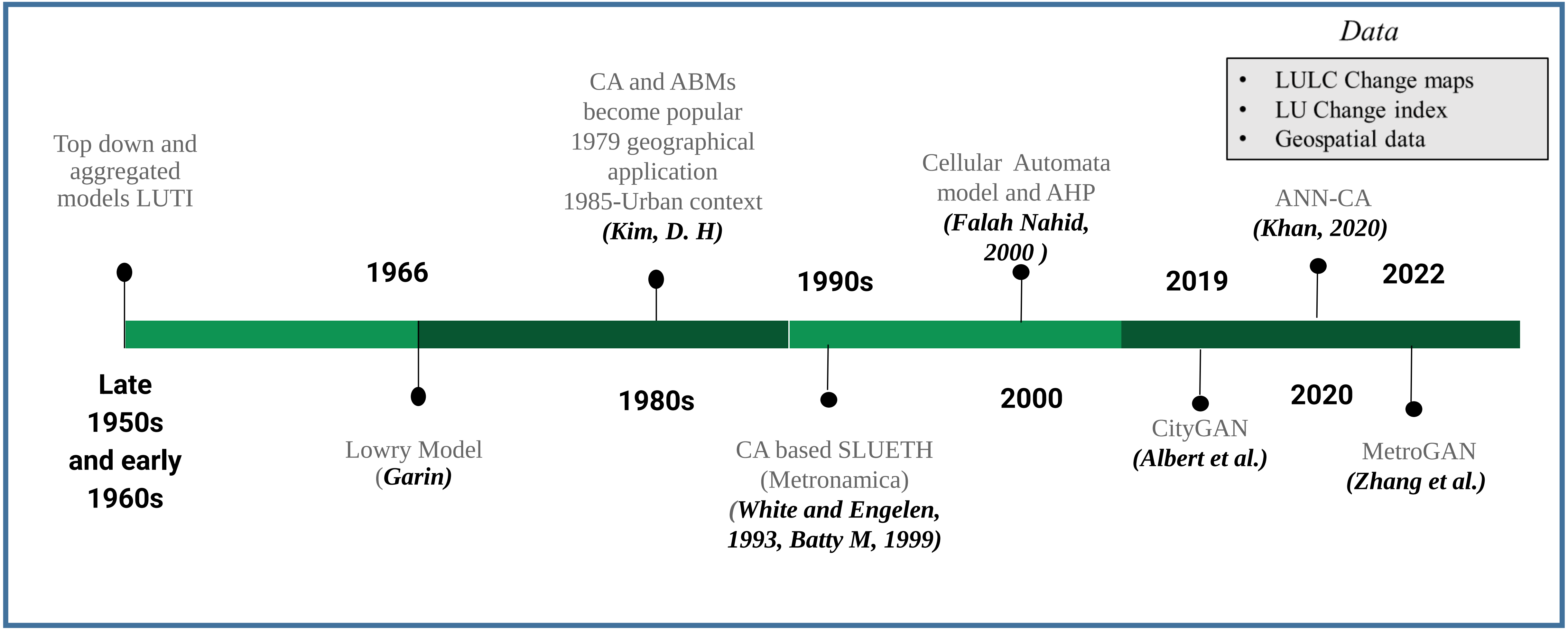
- It also involves determining if the needed infrastructure to support the building is there or needs to be added.

- The data provides context to planners and includes things such as geographic factors, socioeconomic statistics and human mobility.

EO Data in Predicting Urban Settlements

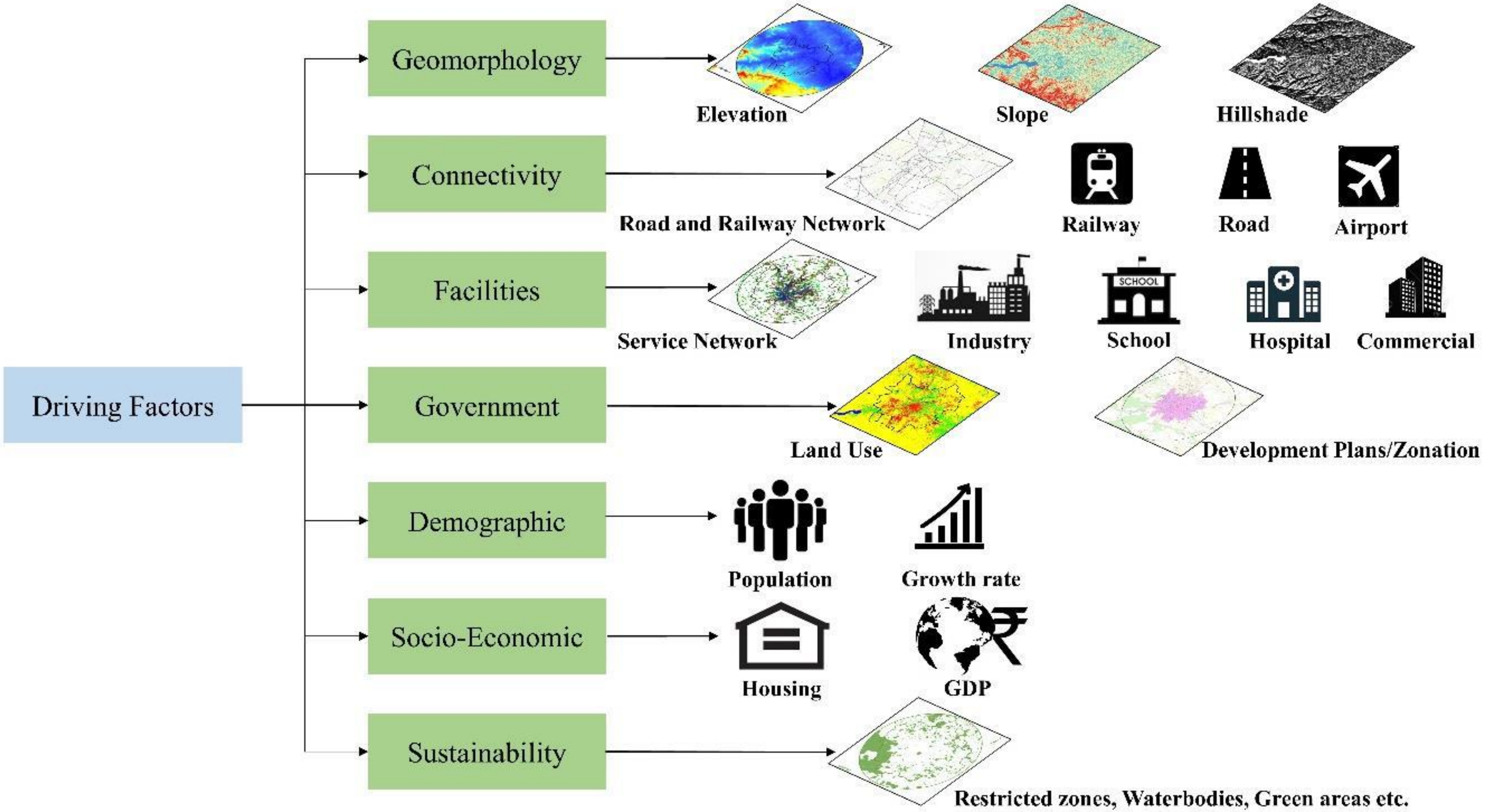


Historical Urban Models



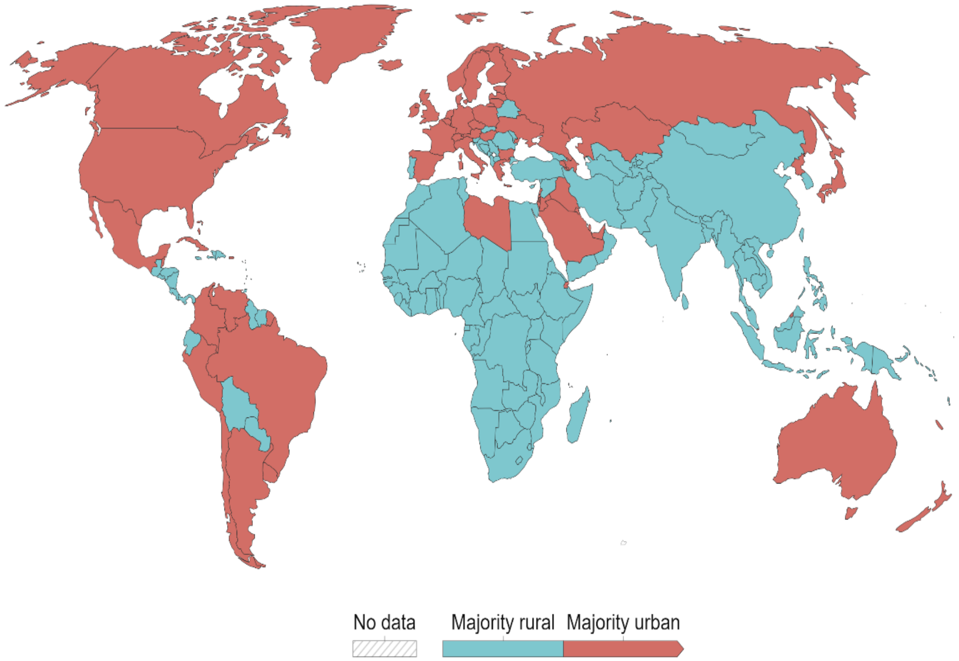
CityGAN is based on a deep Generative Adversarial Network, a machine-learning technique in which a machine is trained in a specific 'set.' This learning occurs after the machine is presented with a data distribution, or a list that includes all points of data; with this information the machine can produce infinite probability points that could be inserted into a graph.

Urban growth driving factors considered by various researchers for developing models

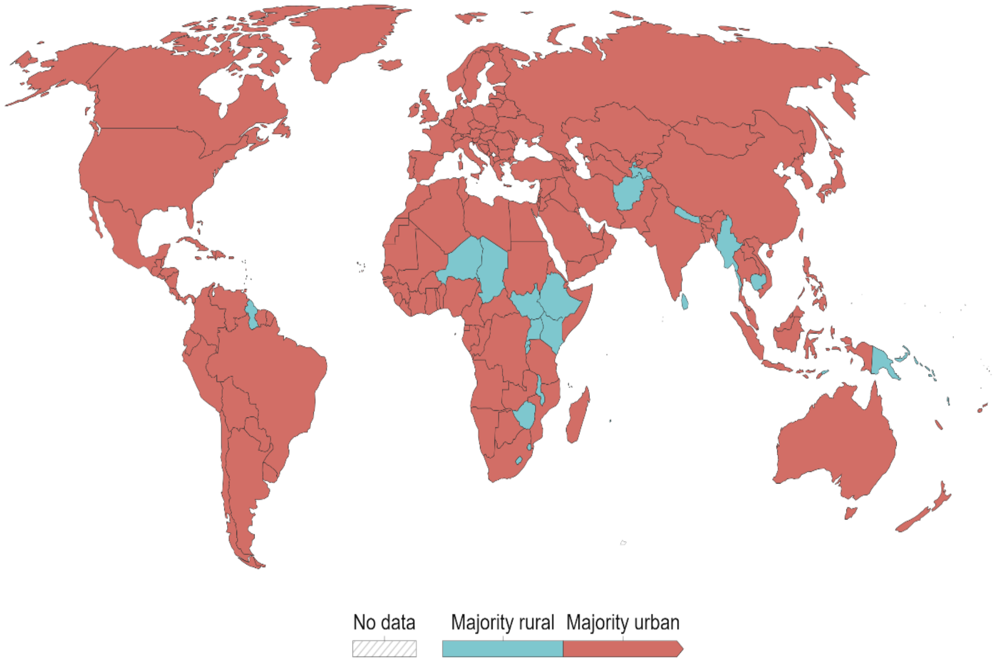


Urbanization, its impact on landscapes

- Today more people live in cities than in rural areas globally.
- Rapid growth of urban areas can be attributed to **two factors**: a natural increase in population (excess of births over deaths); and migration to urban areas.
- 77% of the global urban population (3.26 billion) is expected to be in **developing countries** by 2030.



Source: OWID based on UN World Urbanization Prospects (2018) & Historical Sources (see Sources tab)
OurWorldInData.org/urbanization • CC BY



Source: OWID based on UN World Urbanization Prospects (2018) & Historical Sources (see Sources tab)
OurWorldInData.org/urbanization • CC BY

Population of Cities in India (UN)

Rank	City	Population		Share (%)	Change	
		2020	2021		Population	(%)
1	Delhi	30,290,936	31,181,377	2.24	890,441	2.94
2	Mumbai (Bombay)	20,411,274	20,667,655	1.48	256,381	1.26
3	Kolkata (Calcutta)	14,850,066	14,974,073	1.07	124,007	0.84
4	Bangalore	12,326,532	12,764,935	0.916	438,403	3.56
5	Chennai (Madras)	10,971,108	11,235,018	0.806	263,910	2.41





Motivation: Need for sustainable urban planning

- This rapid demographic and spatial transformation may prove to be difficult for cities in developing countries, especially larger metropolitan and medium-to-small cities, where capacity is typically inadequate to cope with major urban challenges.

- Urban land-use planning, if led by well-informed policies based on sustainable development principles and supported by well-planned and well-managed initiatives can help address many challenges.

- IDSMT Scheme of GoI for small and medium sized cities.

Research Questions?

a. How to generate an urban universe for India based on spatial patterns via learning urban patterns?

b. Is there a relationship between HSI and TI in small and medium cities in India?

c. How to predict (forecast) TI for synthetic (future) urban cities generated by for developing countries like India?

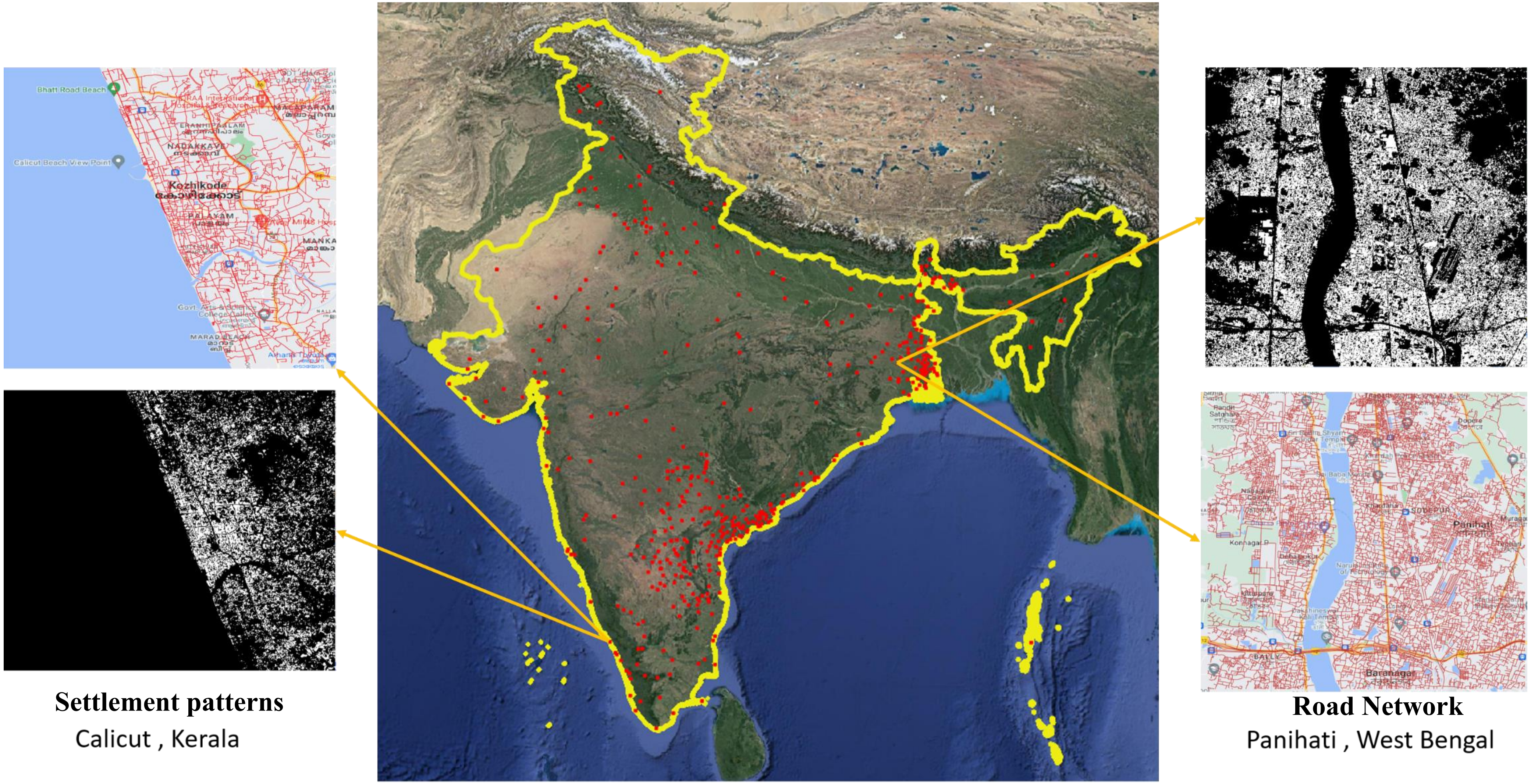
Developing an Integrated Solution

a. To use CityGAN for modelling urban morphology while simulating the spatial structure of small and medium sized Indian cities.

b. To develop a spatial relationship model between derived spatial metrics and topological indices from real city images via Chatterjee Correlation Coefficient.

c. Develop a prediction model to predict the topological index (Network Density) for the settlement patterns.

Study area & data collection



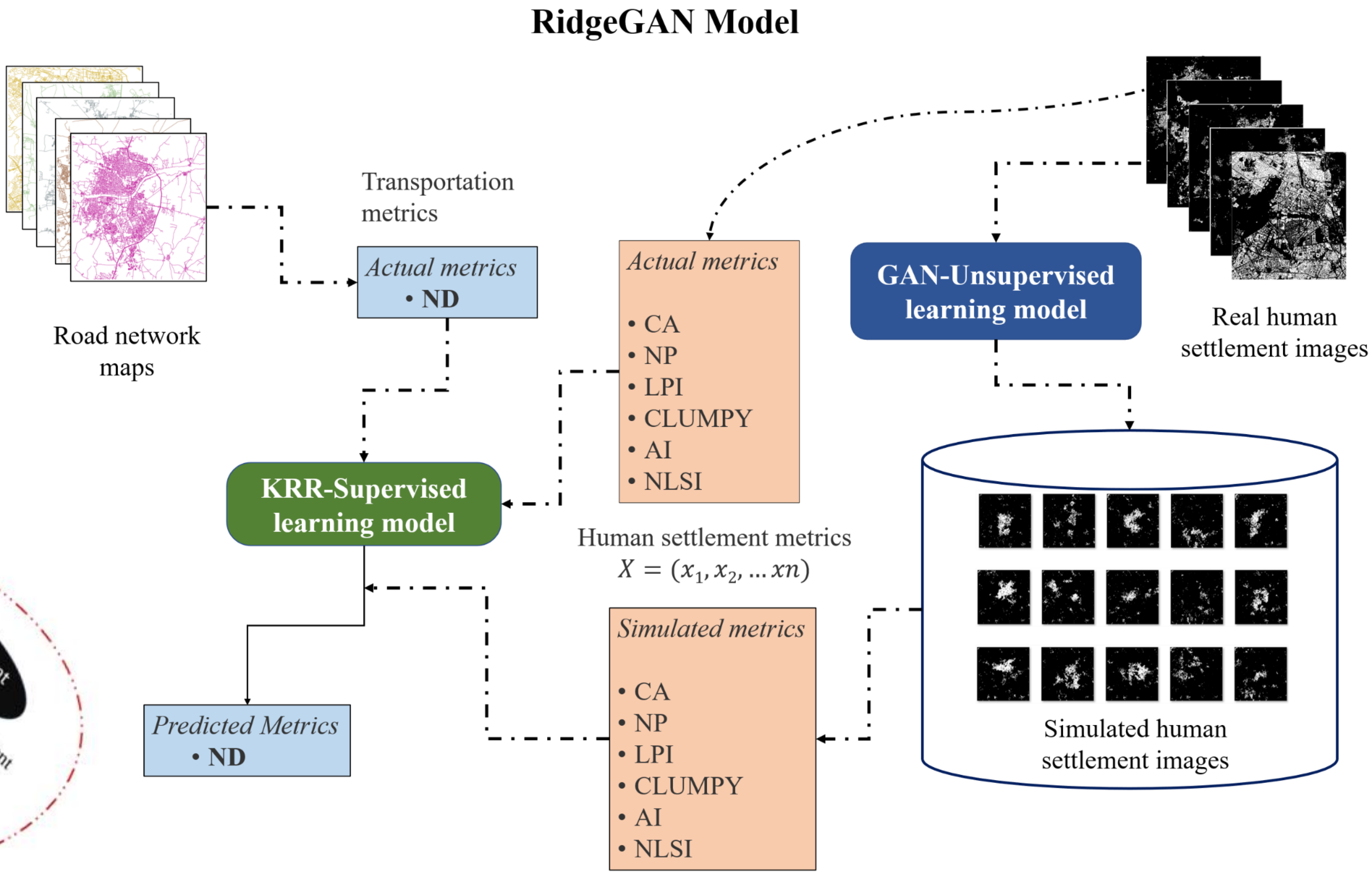
Settlement patterns
Calicut , Kerala

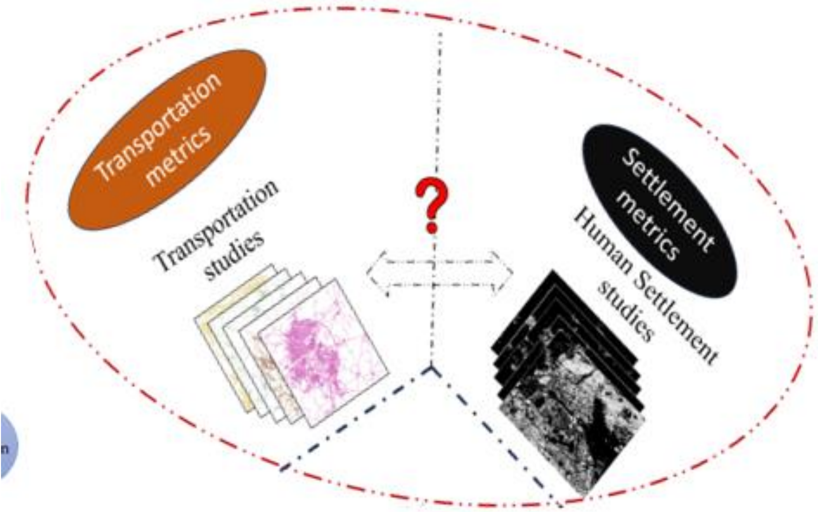
Road Network
Panihati , West Bengal

(a) (b) (c)

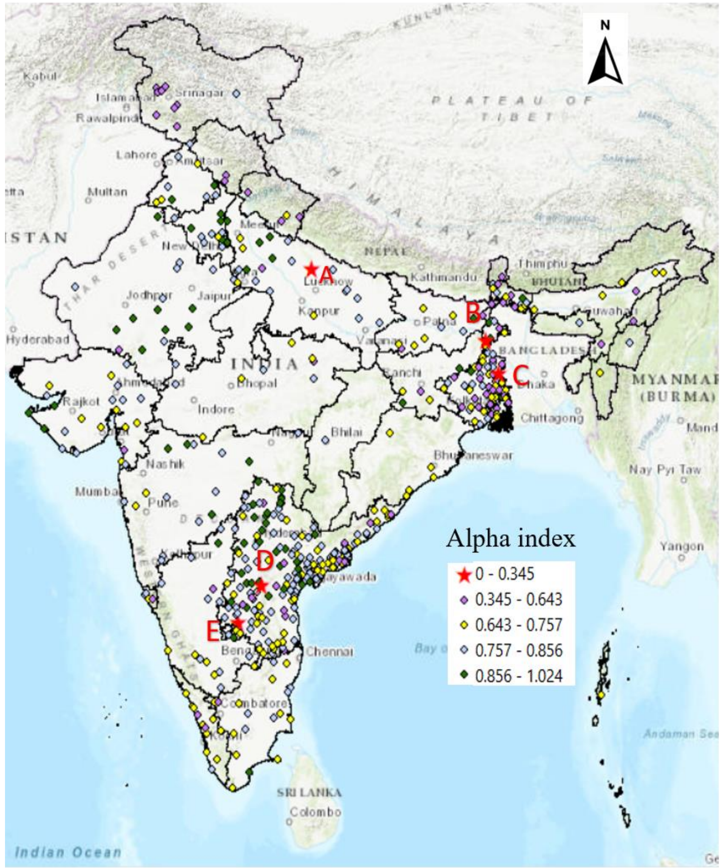
Proposed Solution

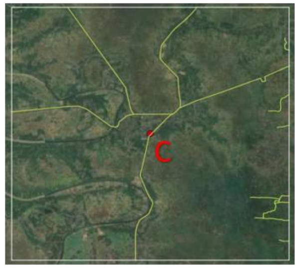
- Implementing CityGAN to generate small and medium-sized Indian cities.
- Assessing the relations between the HSIs and TI, and building a supervised learning model to predict the TI for GAN-generated urban universe.

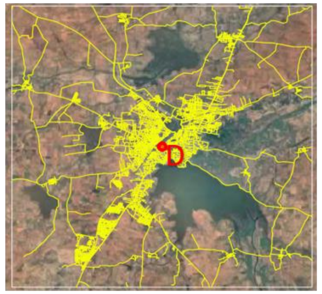




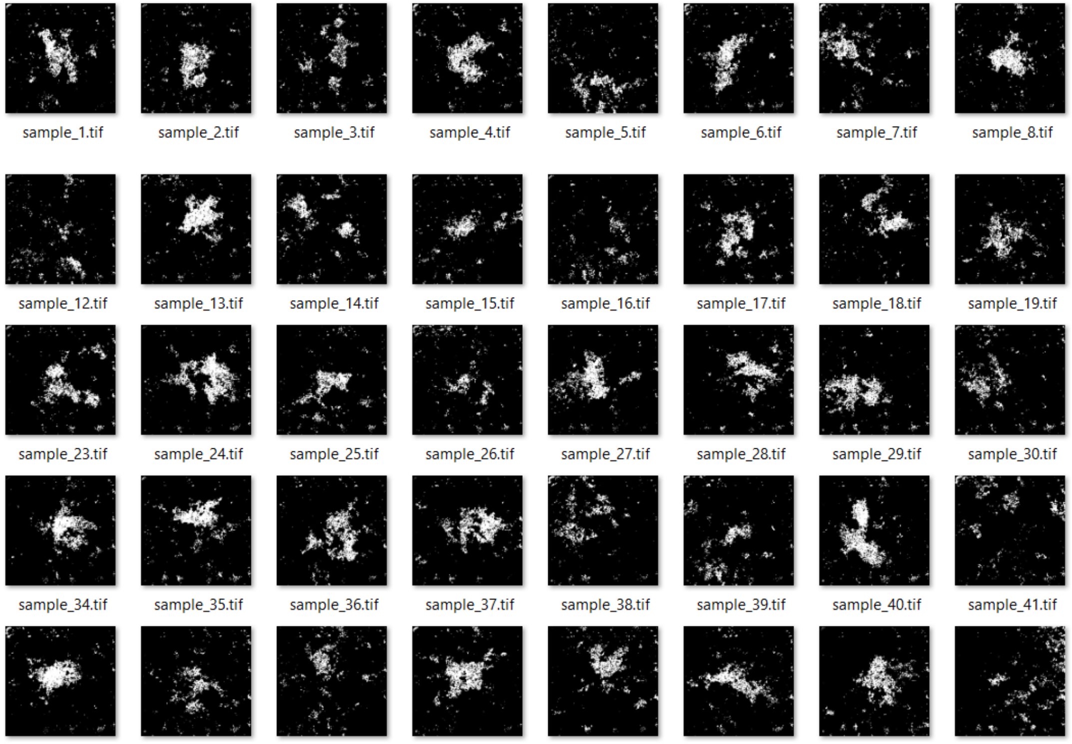
RidgeGAN model - IDSMT Project



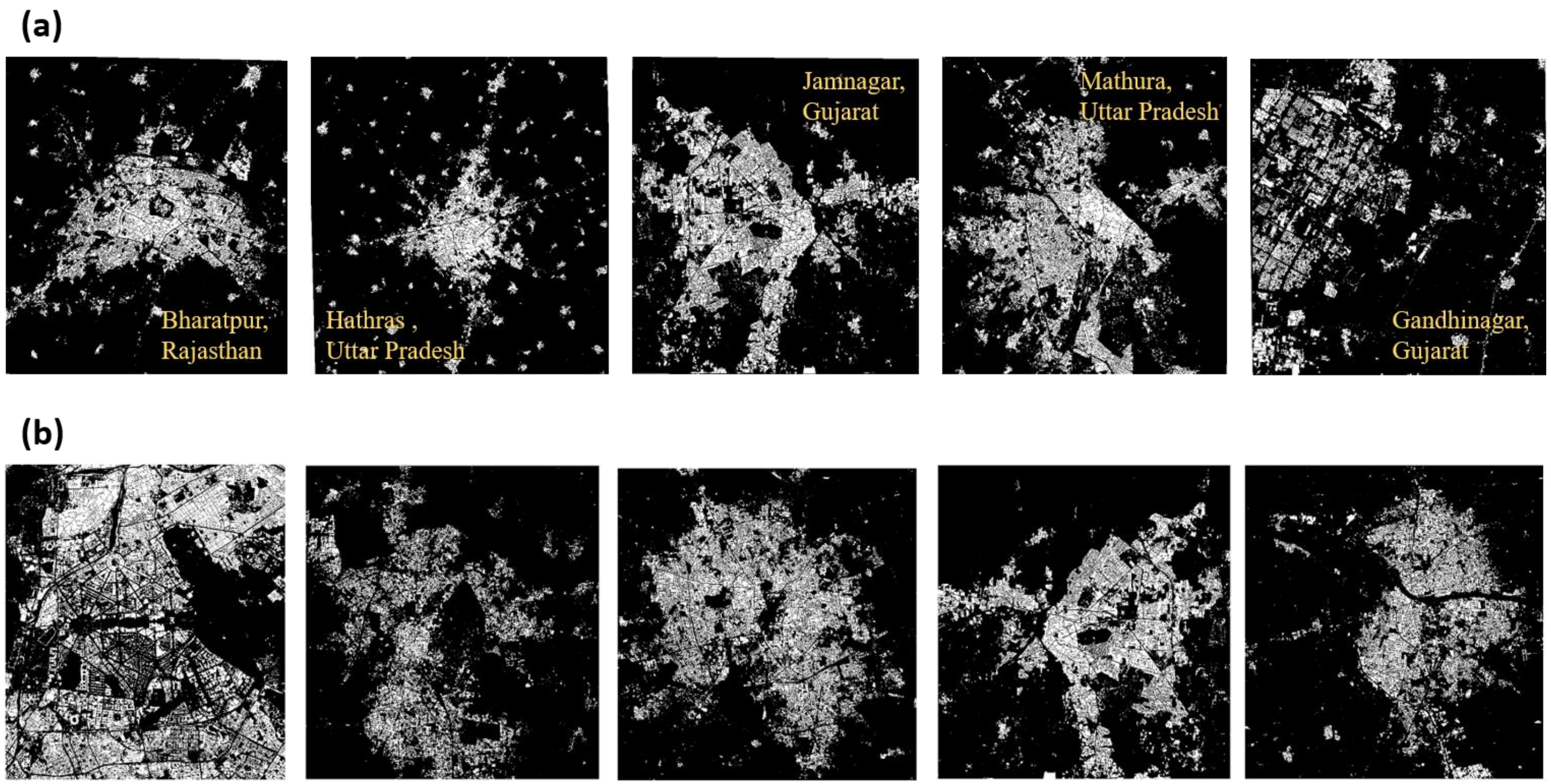


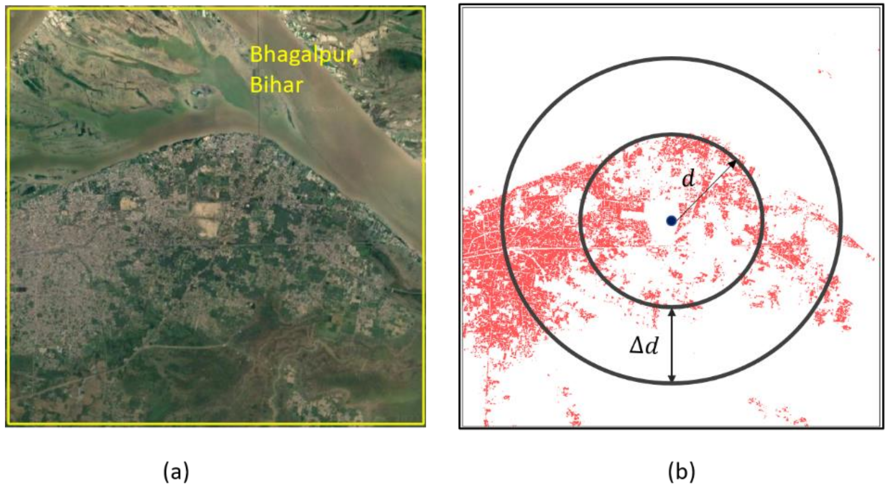


Location	City Name	Population	Alpha	ND
A	Sitalpur	164435	0.343	0.602408
B	Panchanandapur	26358	0.15	0.14
C	Tehata	21093	0.345	0.972
D	Dharmavaram	121874	0.04	4.65
E	Velugodu	23048	0	0



Comparison and Validation





(a) Google satellite view of Bhagalpur city (randomly selected city to explain) in Bihar state (b) human settlement map with the method of computing their average radial profiles.

The pixel values in each case are in the range [0, 1], where 1 represents the portion of land occupied by buildings.

Predicting TI using HSIs:

Performance metrics for the test set of real dataset. Best model's values are highlighted in **bold**.

Accuracy metrics	SVR	DT	GB	MLP	XGB	LR	RF	RR	KRR
MSE	6.439	5.792	5.265	4.505	4.448	4.343	4.133	3.998	**3.661**
MAE	1.832	1.716	1.688	1.582	1.535	1.576	1.475	1.511	**1.397**
R^2 Score	0.490	0.541	0.583	0.643	0.648	0.656	0.673	0.683	**0.710**
Adj R^2 Score	0.463	0.517	0.561	0.624	0.629	0.638	0.655	0.667	**0.695**

Limitations

- ARP and peak search algorithm evaluation methods particularly vary based on the **quality and diversity** of the generated synthetic data. They may fail when it comes to time-varying urban patterns and enormous data for analysis.

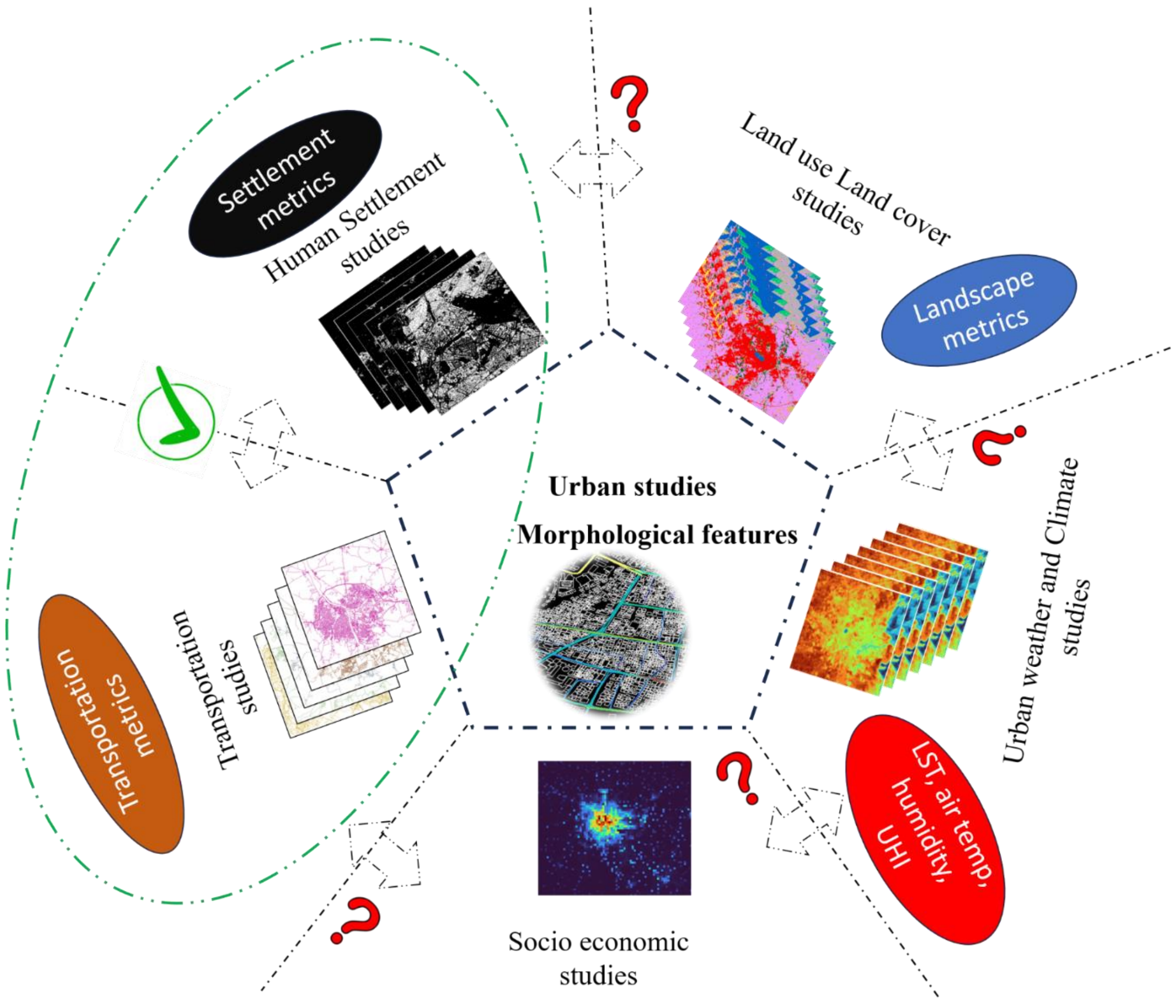
- Although this study evaluates the relationship between HSIs and TI based on OSM data, this validation can not be performed on the simulated data for which the network density is not available.

Our Team and Ongoing Research

Rahisha Thottolil
PhD Student, IIITB

Uttam Kumar
Faculty, IIITB

Tanujit Chakraborty
Faculty, Sorbonne



Appendix-1: Generative adversarial network (GAN)

- Generative models are capable of making new plausible data.
- The term "adversarial" in this case pertains to a unique architecture for training an effective generator network.
- Random input vectors are given to the generator to make plausible samples that are ideally discriminated 50:50 in a fully trained model. However, some issue in training GANs include:

- Vanishing gradient
  - Best if discriminator starts in a less robust state so it is improving along with the generator rather than being too sophisticated from the start
  - Modified minimax loss can help with this, proposed in original paper, maximizing log(D(G(z))) instead of minimizing it.
- Mode Collapse
  - Local minimum in discriminator training, Wasserstein loss can help with this, uses critic model, D(x) – D(G(z)), rather than a threshold valued discriminator
- Failure to Converge
  - Regularization of discriminator can help with this
  - Over training generator past random feedback from discriminator must be monitored

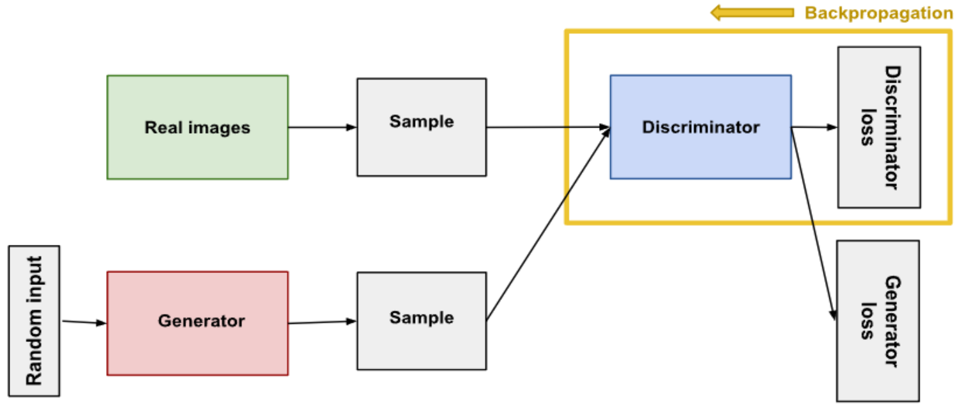


Figure 1: Backpropagation in discriminator training.

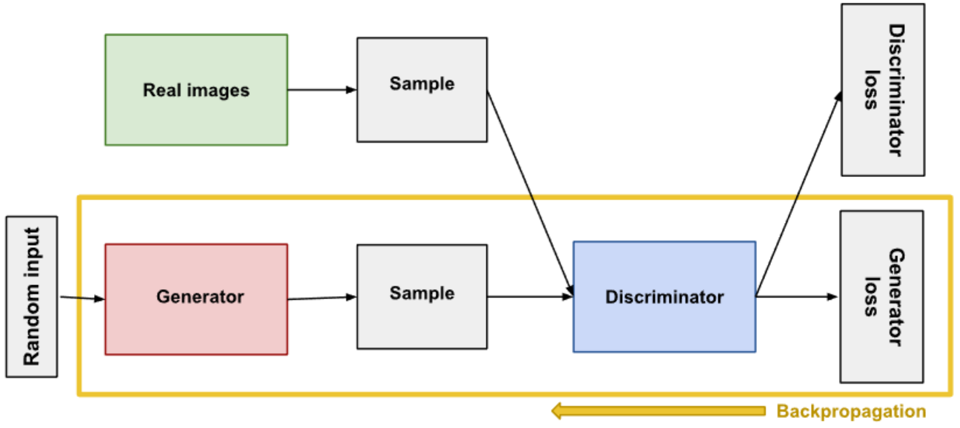


Figure 1: Backpropagation in generator training.

https://developers.google.com/machine-learning/gan/discriminator

Generative Adversarial Networks. Goodfellow 2014

Appendix-2: Kernelized Ridge Regression

- Recall the ridge regression problem: $$\boldsymbol{w} = \arg\min_{\boldsymbol{w}} \sum_{n=1}^{N} (y_n - \boldsymbol{w}^\top \boldsymbol{x}_n)^2 + \lambda \boldsymbol{w}^\top \boldsymbol{w}$$
- The solution to this problem was

$$\boldsymbol{w} = (\sum_{n=1}^{N} \boldsymbol{x}_n \boldsymbol{x}_n^\top + \lambda \mathbf{I}_D)(\sum_{n=1}^{N} y_n \boldsymbol{x}_n) = (\mathbf{X}^\top \mathbf{X} + \lambda \mathbf{I}_D)^{-1} \mathbf{X}^\top \boldsymbol{y}$$

- Can use matrix inversion lemma: $$(\mathbf{F}\mathbf{H}^{-1}\mathbf{G} - \mathbf{E})^{-1}\mathbf{F}\mathbf{H}^{-1} = \mathbf{E}^{-1}\mathbf{F}(\mathbf{G}\mathbf{E}^{-1}\mathbf{F} - \mathbf{H})^{-1}$$
- Using the lemma, can rewrite $\boldsymbol{w}$ as

$$\boldsymbol{w} = \mathbf{X}^\top (\mathbf{X}\mathbf{X}^\top + \lambda \mathbf{I}_N)^{-1} \boldsymbol{y} = \mathbf{X}^\top \boldsymbol{\alpha} = \sum_{n=1}^{N} \alpha_n \boldsymbol{x}_n$$ where $$\boldsymbol{\alpha} = (\mathbf{X}\mathbf{X}^\top + \lambda \mathbf{I}_N)^{-1} \boldsymbol{y} = (\mathbf{K} + \lambda \mathbf{I}_N)^{-1} \boldsymbol{y}$$

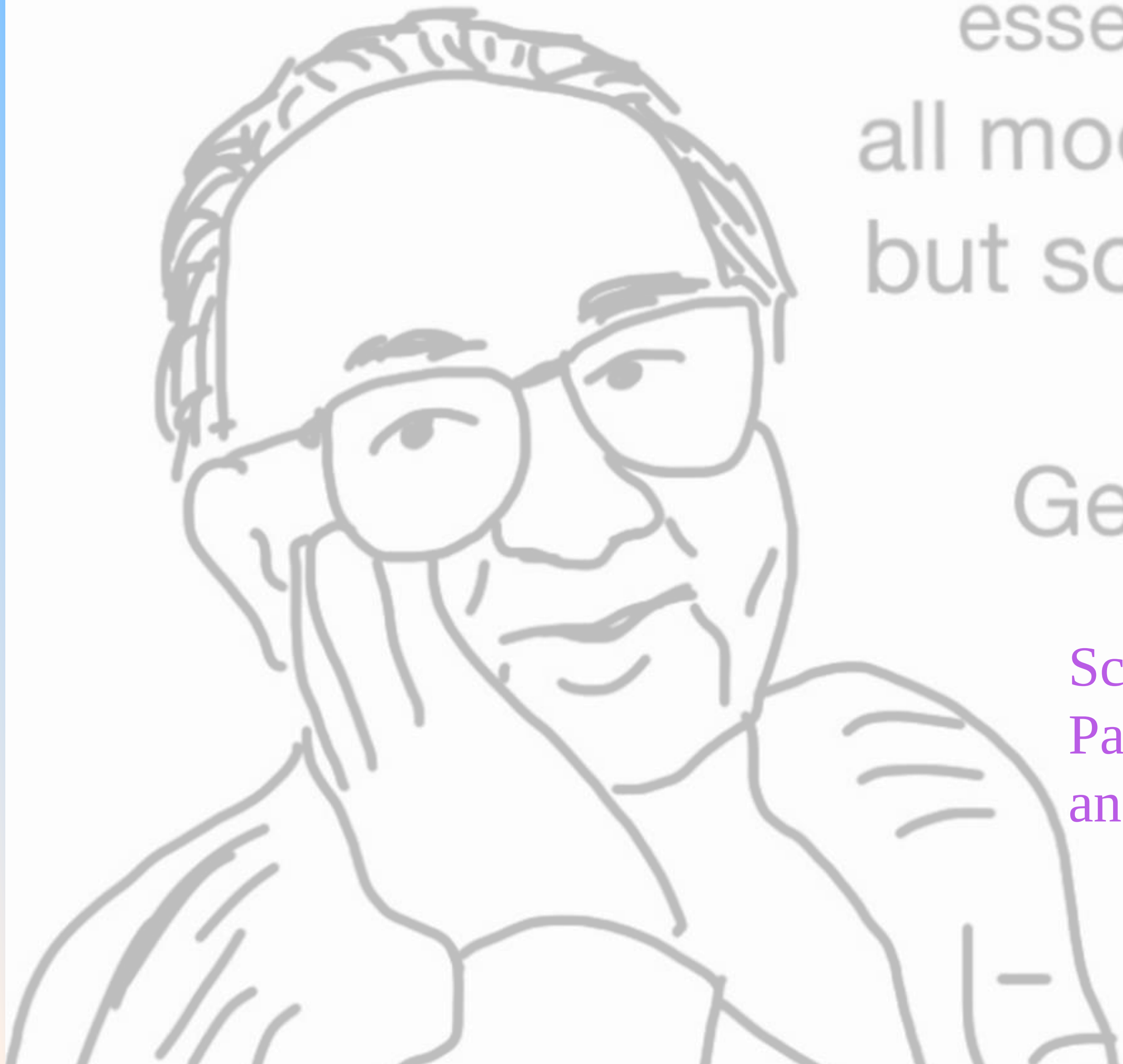
$N \times 1$ vector of dual variables

Note: Not sparse unlike SVM

- Kernelized weight vector will be $\boldsymbol{w} = \sum_{n=1}^{N} \alpha_n \phi(\boldsymbol{x}_n)$
- Prediction for a test input $\boldsymbol{x}_*$ will be $\boldsymbol{w}^\top \phi(\boldsymbol{x}_*)$ $$= \sum_{n=1}^{N} \alpha_n \phi(\boldsymbol{x}_n)^\top \phi(\boldsymbol{x}_*) = \sum_{n=1}^{N} \alpha_n k(\boldsymbol{x}_n, \boldsymbol{x}_*)$$

Prediction cost is also linear in N (like KNN)

THANK YOU!



essentially,
all models are wrong,
but some are useful

George E. P. Box

Scan me for Paper, Data, and Code: