HPC in the social sciences

How does Strato, uCloud and LUMI make a difference in the social sciences?

The things I do

Background

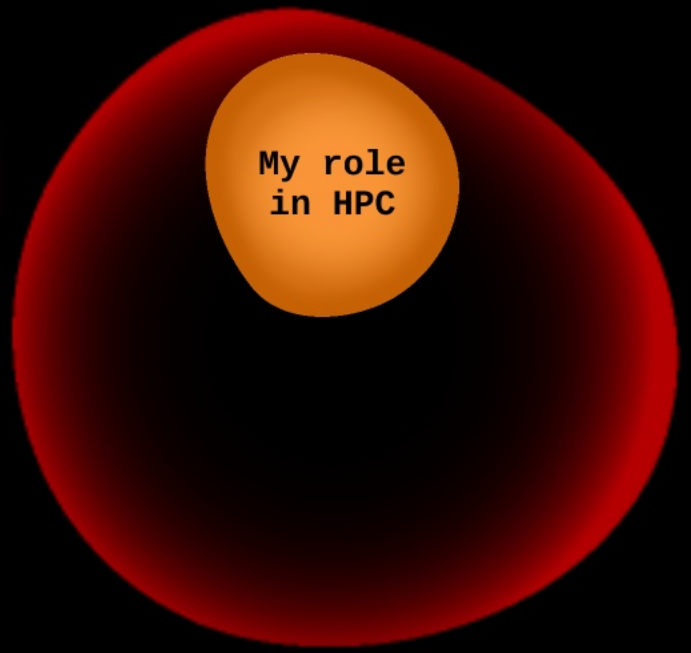
Teaching

Research

What I do

I am currently

- Using transformers to help doctors figure out exactly how much pee is too much pee
- Using satellite images to figure out what neighborhoods look like and what that means
- Checking if full moon has anything to do with suicide rates and if Zodiac signs have any effect on... anything
- Building a model to compare different types of supervised algorithms to predict cardiovascular disease
- Running unsupervised models to figure out if there are any regional logic in where health services are located after centralization
- Using spatially weighted clustering to investigate if gambling machines are being placed in a neighborhood because it is poor or if the neighborhood becomes worse off when gambling machines enter the chat
- Developing a method for bias testing algorithms used in the public sector in risk scores
- Doing transfer learning to change an algorithm designed for mold detection in organic material to look for unkempt gardens



My role in SSH

In regards to the use of data processig machines

Scientific co-director of the center for computational SSH

Coordinator for a spezialized masters program in social data science

Working with computationally heavy machinery in a wide array of research grants ranging from smaller national grants to horizon projects

HPC in the social sciences

How does Strato, uCloud and LUMI make a difference in the social sciences?

The things I do

Background

Teaching

Research

Core concept

What defines a neighborhood?

Most of us readily use some type of administrative division to "simplify" our knowledge about the physical world

We (to some extent) identify with our country, we live part of that country, we know some of the differences between municipalities, we know the city we live in and, to some extent feel like part of a much smaller part of that city - our neighborhood

What are neighborhoods?

Why should we care about neighborhoods?

A montage

Should we care?

Neighborhoods play a huge part of not only our daily lives but also as something that affects us while living there and even after moving - not in the sense that they are toxic but in the sense that neighborhood dynamics help shape who we are and who we become

HPC in the social sciences

How does Strato, uCloud and LUMI make a difference in the social sciences?

The things I do

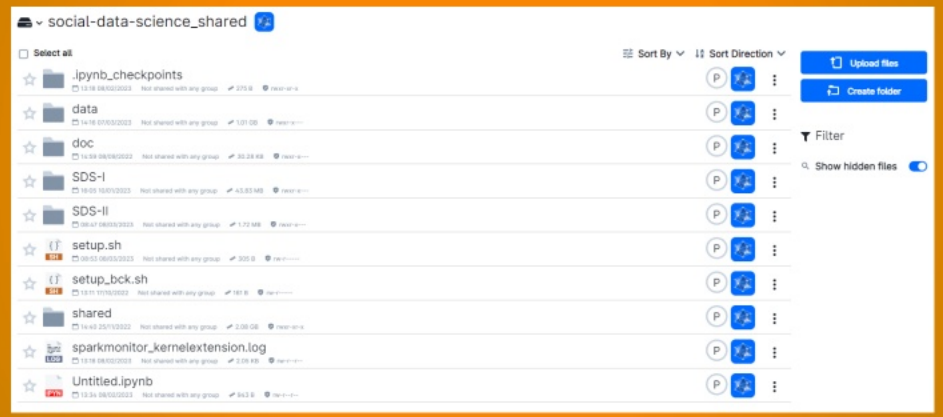
Background

Teaching

Research

Teaching

Social data science in sociology



- Course outline
- why uCloud is indispensable

What we teach

Data wrangling with Python; numbers, text and images

Data scraping and modelling; scrapers, crawlers and API-calls

Data visualizations; relational, dynamic and static data

Clustering and dimensionality reduction

Anomaly detection and association mining

Topic modelleing, word embeddings and language models

Supervised prediction; regression, random forest, neural networks and deep learning

Explainability and "dashboarding" of data

HPC in the social sciences

How does Strato, uCloud and LUMI make a difference in the social sciences?

The things I do

Background

Teaching

Research

But... satellites?

It's all fine and dandy with neighborhoods but what does this have to do with satellites?

We know that the social element of neighborhoods are important but so are the way they look and perhaps we can learn from visuals as well

What are the data/methods?

What are the possibilities?

Data and methods

Data consists of Swedish and Danish register data and a still expanding collection of satellite imagery (mainly from Landsat 8)

The interesting part about this is that we can use a plethora of different indicators to help train different models - it's not a static data set of registers but dynamic

Satellites

Methods

What does it look like?

What does it look like?

What does it look like?

Methods?

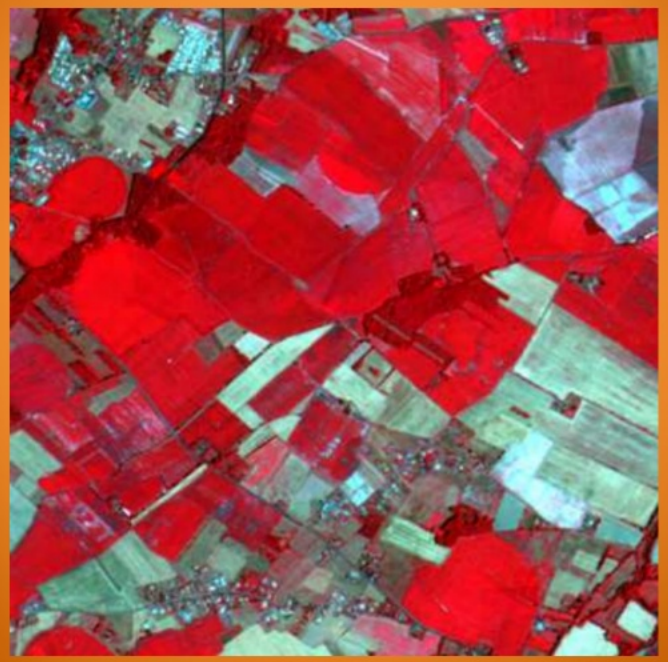
Redefining neighborhoods are mostly done by looking at data but now we have another layer as well; a visual one

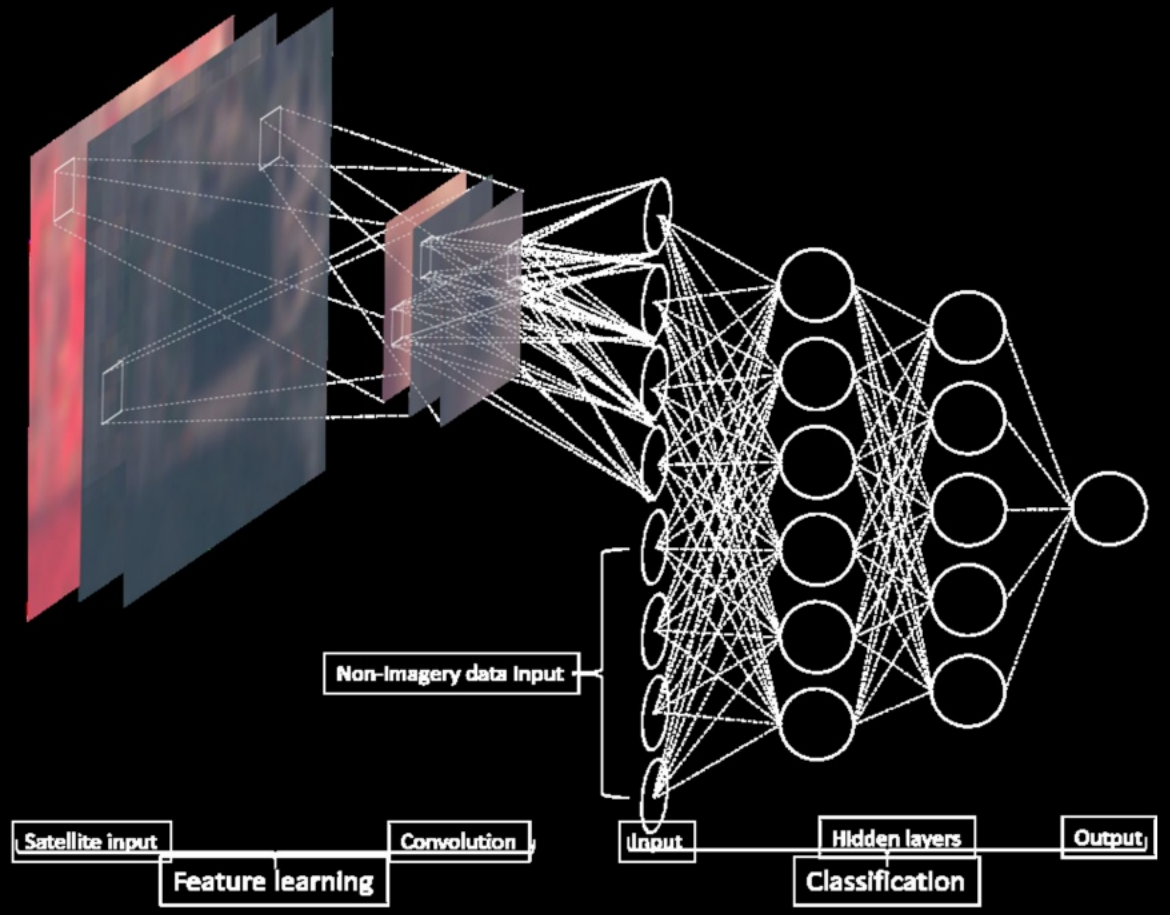
Visual data requires (especially when we have more than 289 million local images) computational help

- Since most common pre-trained image classification models like VGG-16, ResNet50, Inceptionv3 or EfficientNet rely on common problems like finding bikes, cars or apples and this goes somewhat beyond that, we are building our own stuff

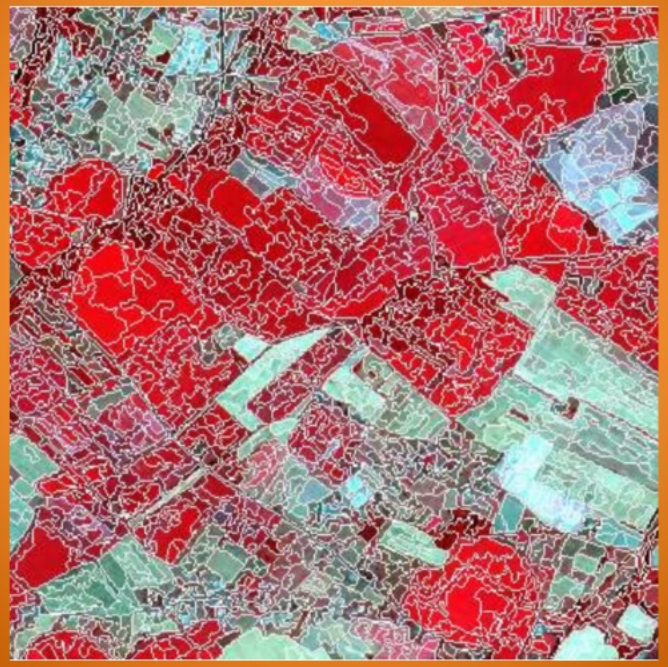
Preliminary solutions

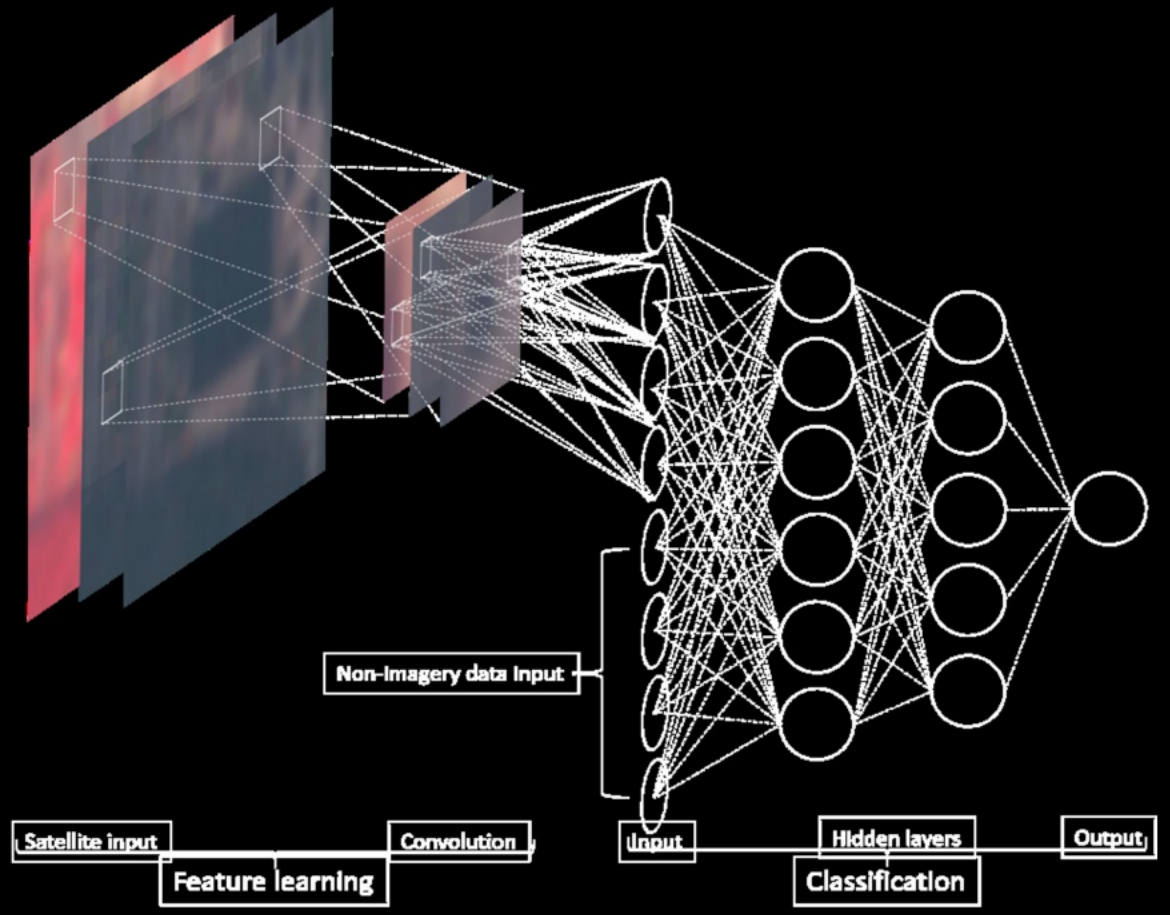
Architecture



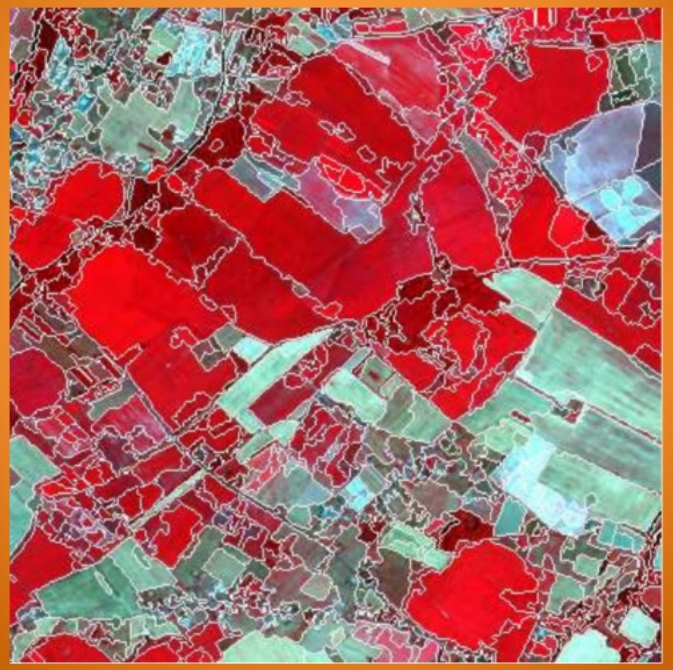


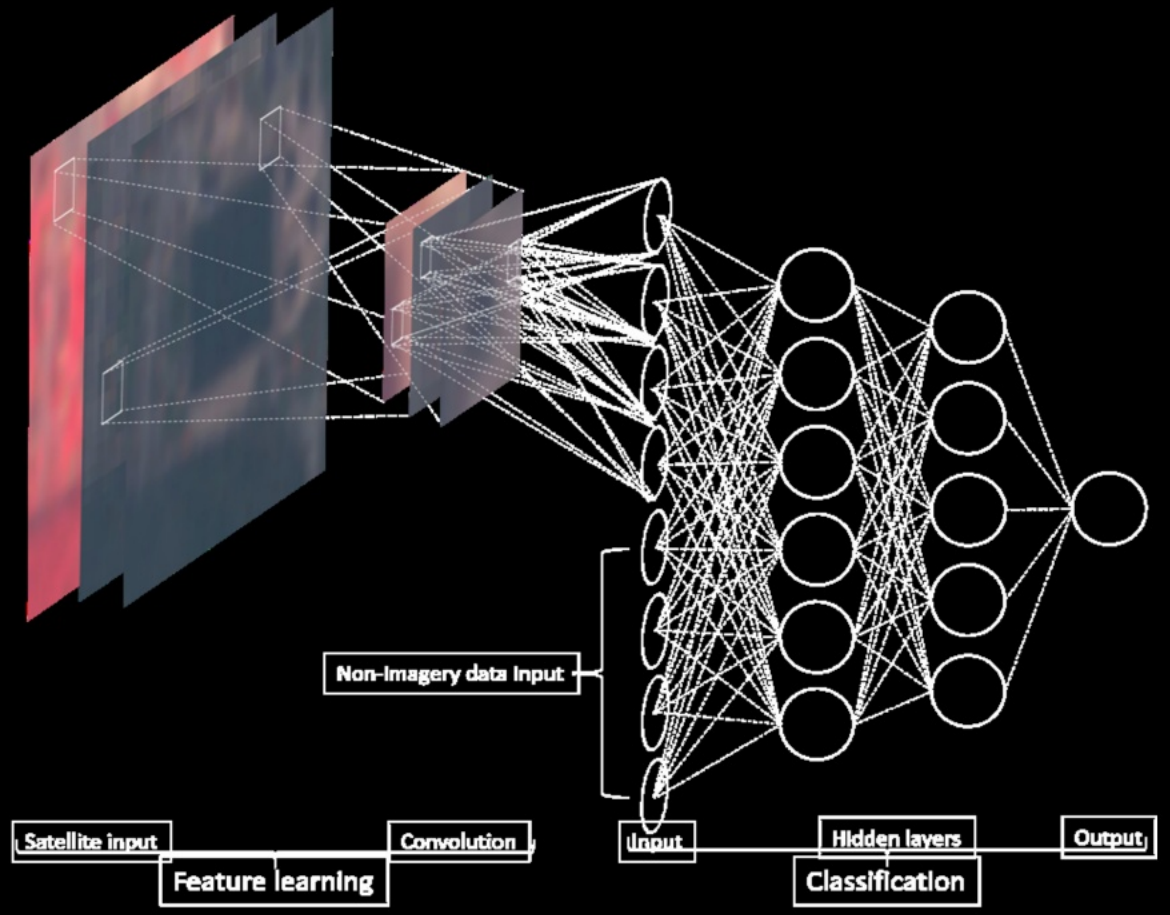
Architecture





Architecture





Possibilities?

Much like personalized medicine, we work with a version of "personalized sociology"

By deconstructing administrative areas and, much more inductively, let the social define the geography we hope to be able to isolate specific types of neighborhoods or isolated enclaves to further how we look at neighborhoods

HPC in the social sciences

How does Strato, uCloud and LUMI make a difference in the social sciences?

The things I do

Background

Teaching

Research