



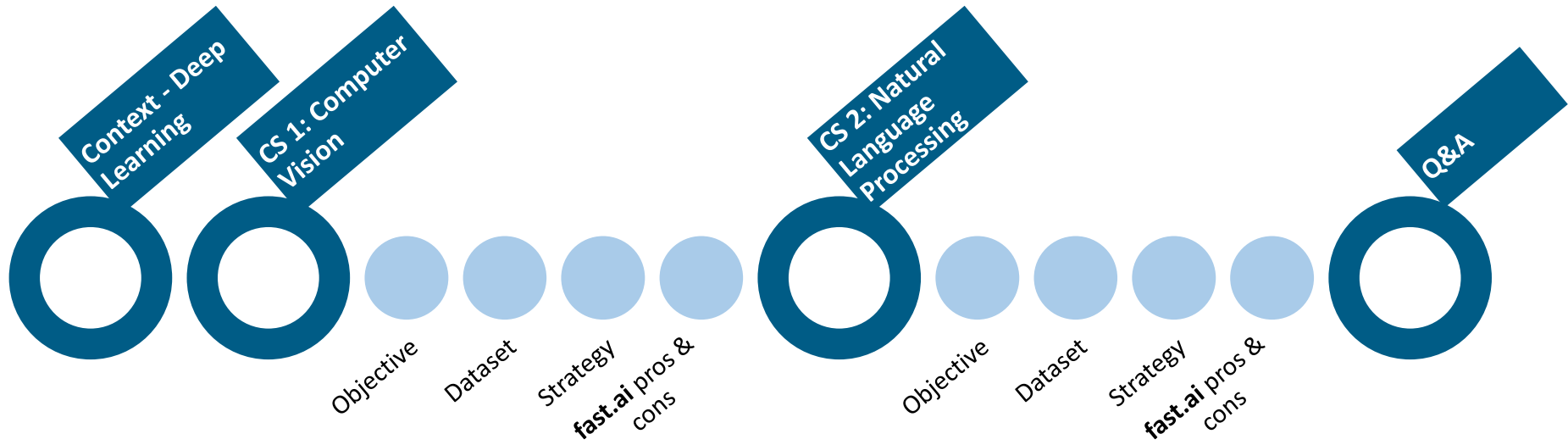
Food and Agriculture Organization
of the United Nations

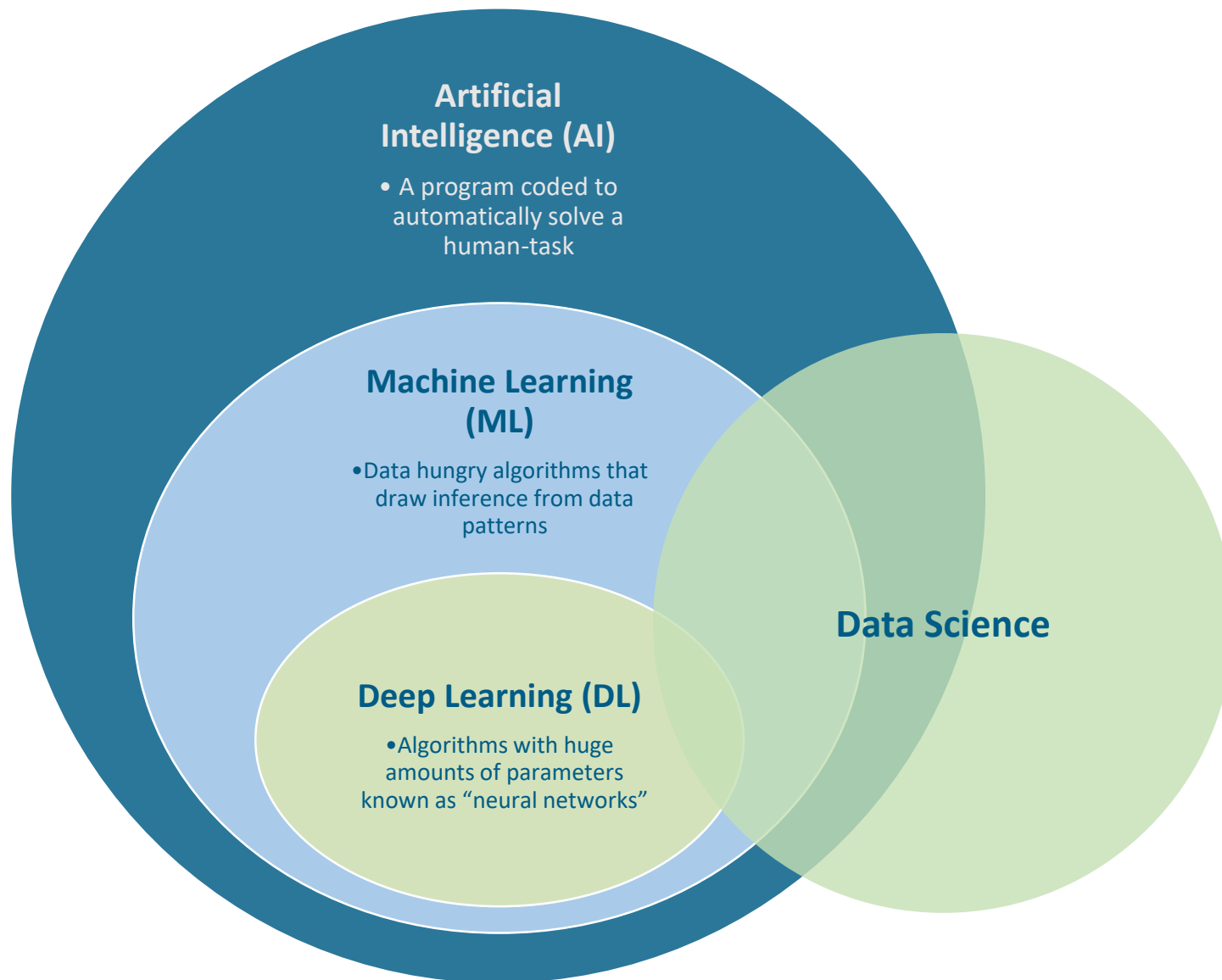
>> FAO Statistics Division

fast.ai: Applications and Lessons Learned

Gianfausto Bottini– Hannah Gerits
15 December 2022, ESS

>> FAO Statistics Division





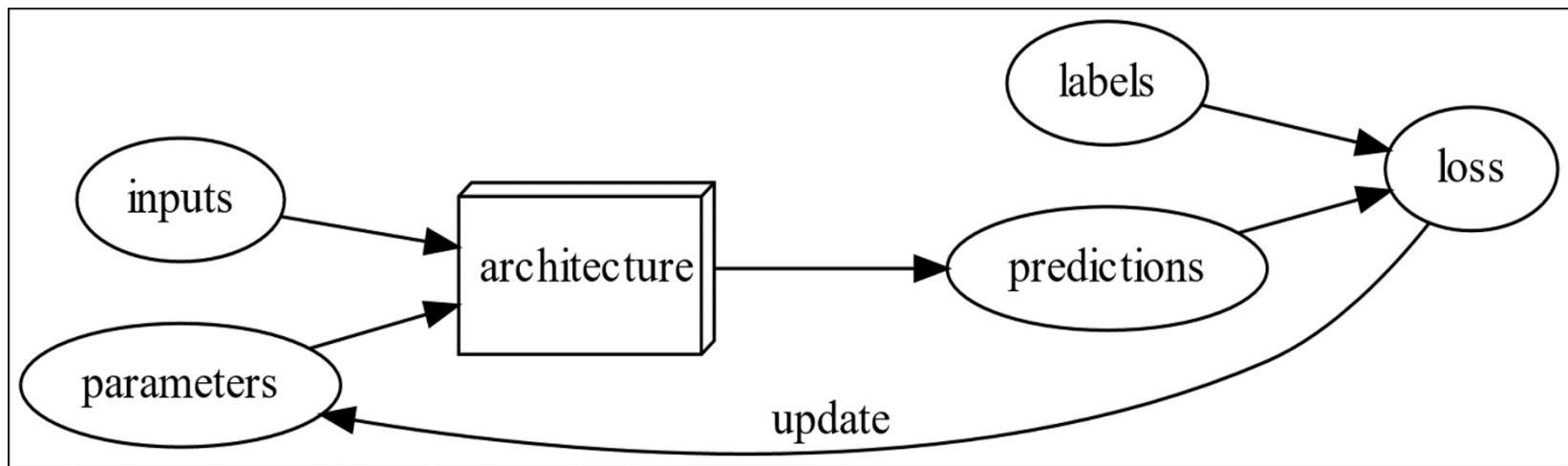
Neural networks in a nutshell

JUST Mathematical functions trained to learn from the past!

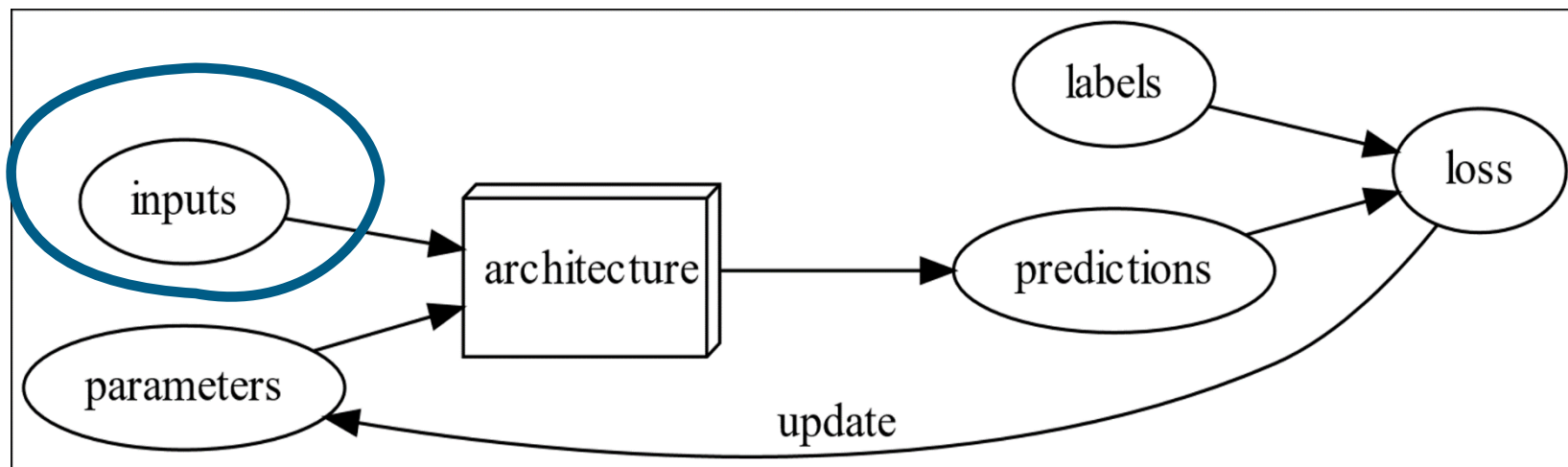
Very good at doing **JUST** one task!



Training a Neural Network: key concepts

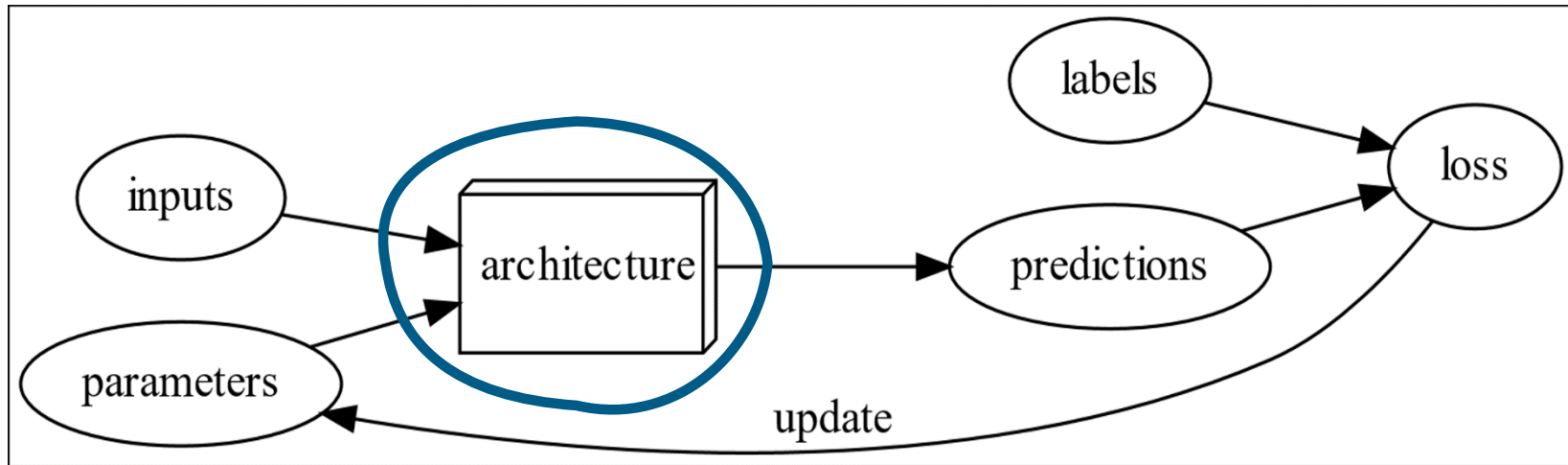


Training a Neural Network: key concepts



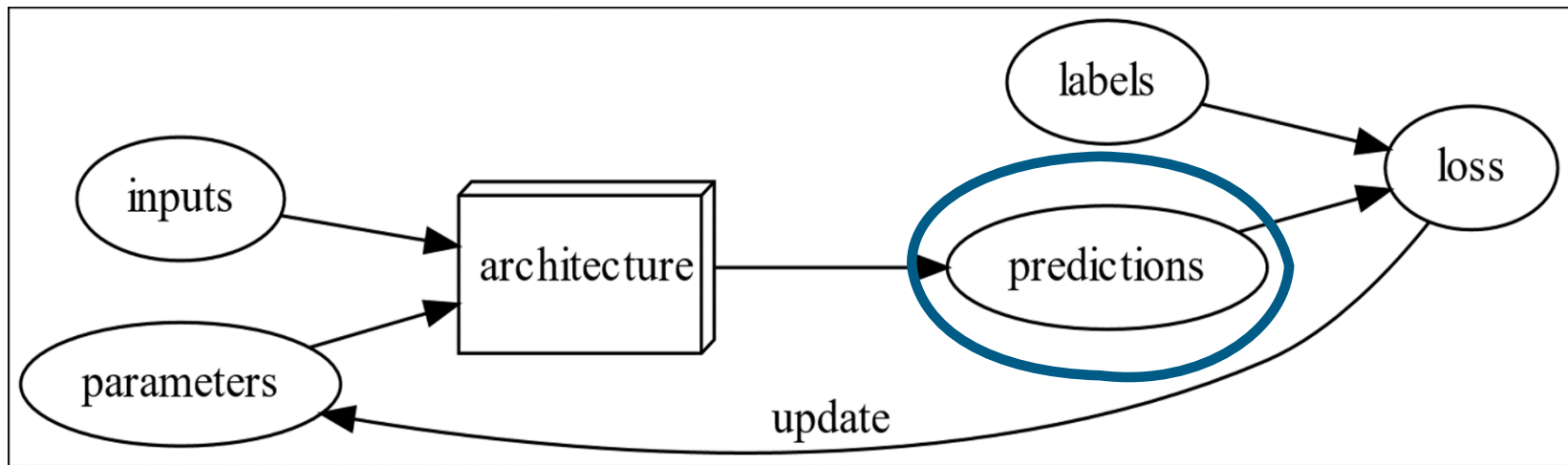
Inputs (or Dataset, or Ground-Truth): a stack of hundreds or thousands of labeled data and can be images, texts, videos etc.

Training a Neural Network: key concepts



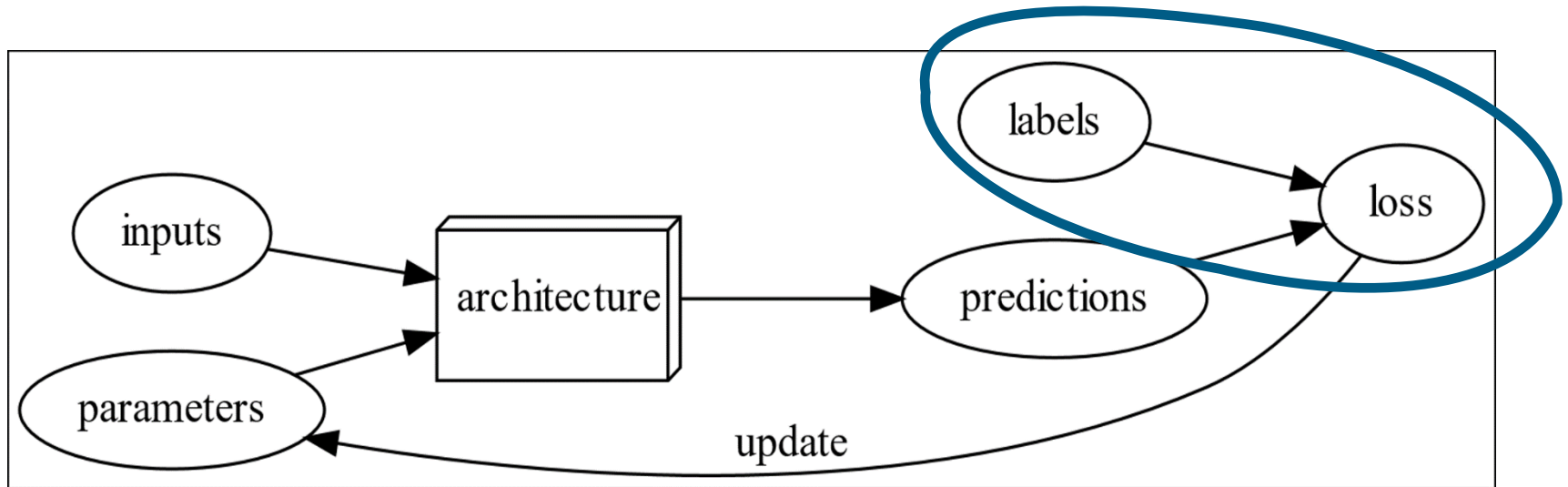
The functional form of the *model* is called its *architecture*. The Data Scientist is able to select the best architecture according to the target and dataset numerosity and quality

Training a Neural Network: key concepts



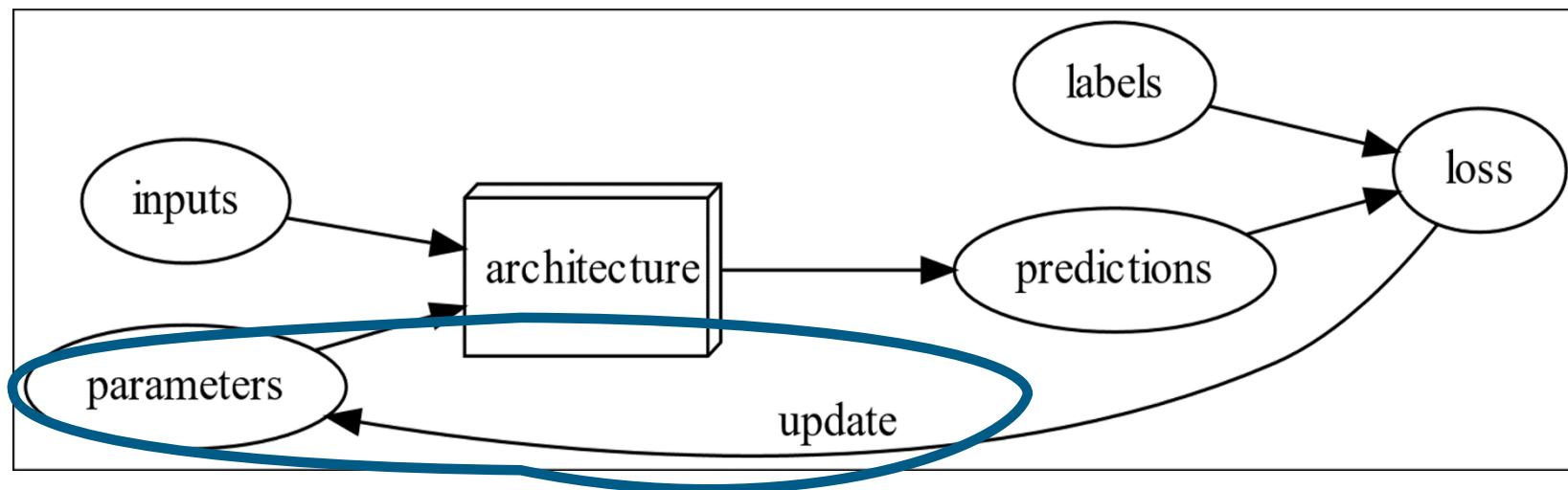
Once the model is trained, also after one or a few epochs, it is able to make predictions on the **validation set**.

Training a Neural Network: key concepts



- Labels and predictions are used to warn the Data Scientist about the current accuracy of the model
- The measure of *performance* is called the *loss*.

Training a Neural Network: key concepts



- After every epoch the parameters (or weights) are updated

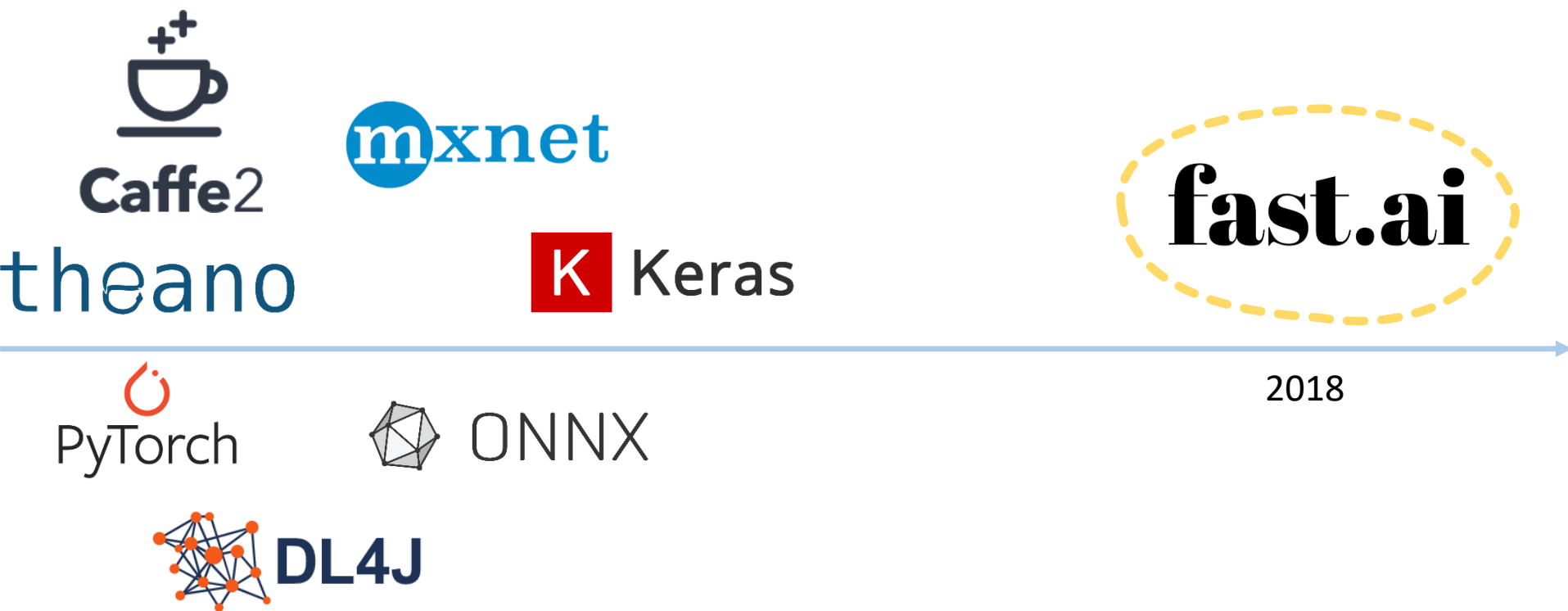
Areas in which DL is now the best



DL Frameworks timeline



DL Frameworks timeline



What is fast.ai?

- Newborn **python (or R) library** for **Deep Learning** projects
- Perfect to build **neural nets (NN)** from scratch or use already implemented ones
- Unlike other frameworks, it allows building NN **in very little time**
- Fast coding, fast training, and fast deployment
- Perfectly embeddable in **Dash**, to create outstanding AI-powered Apps
- Natively working with **Jupyter**
- Full of **embedded functions** helping data visualization
- Tons of **free open-source material**, also including the book in the form of Jupyter notebook pages

Case scenario 1 – Computer Vision (applied to Earth Observation Domain)

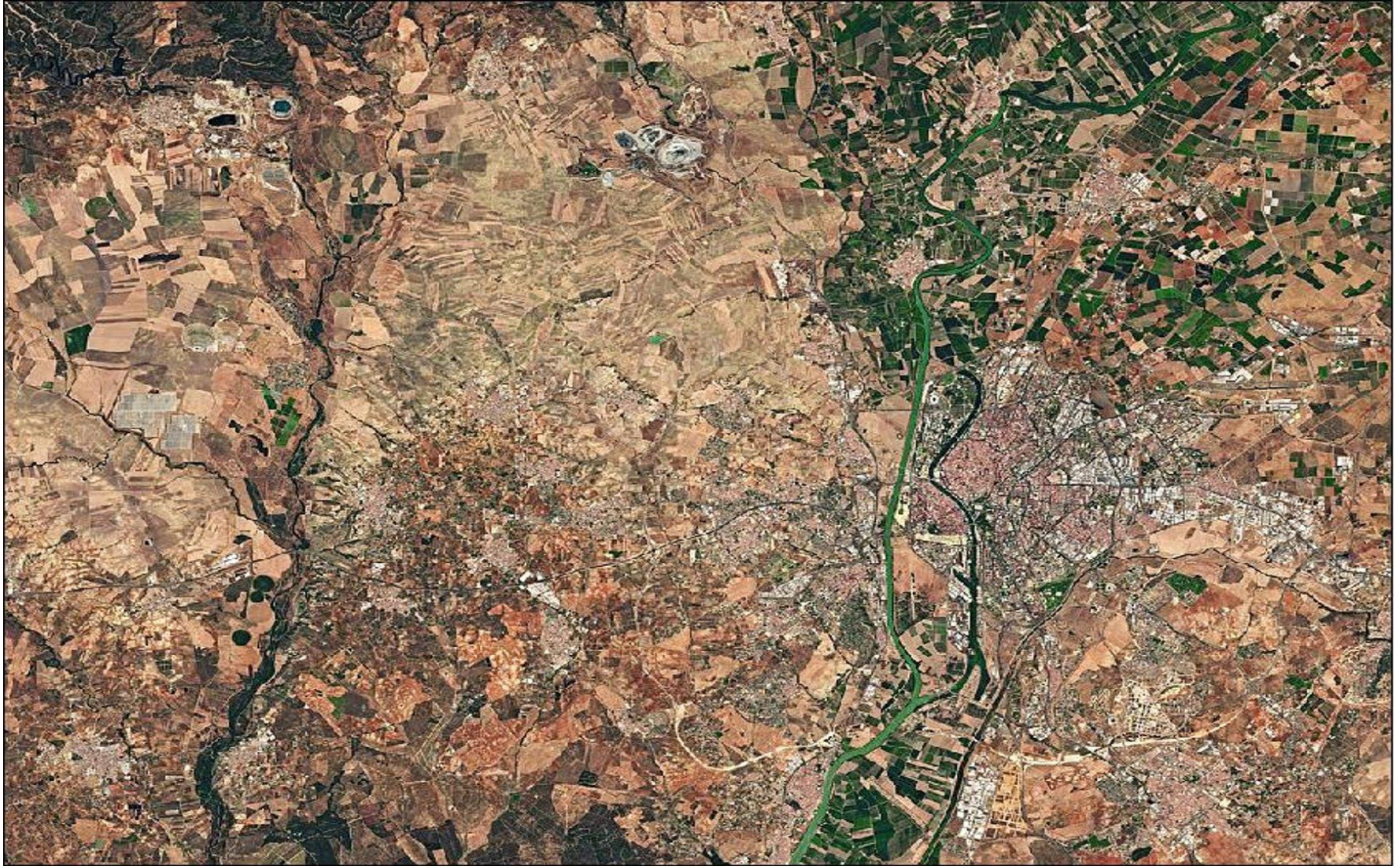
*Give a **coarse estimation of the land cover of countries where little-to-none in-situ data are available.***

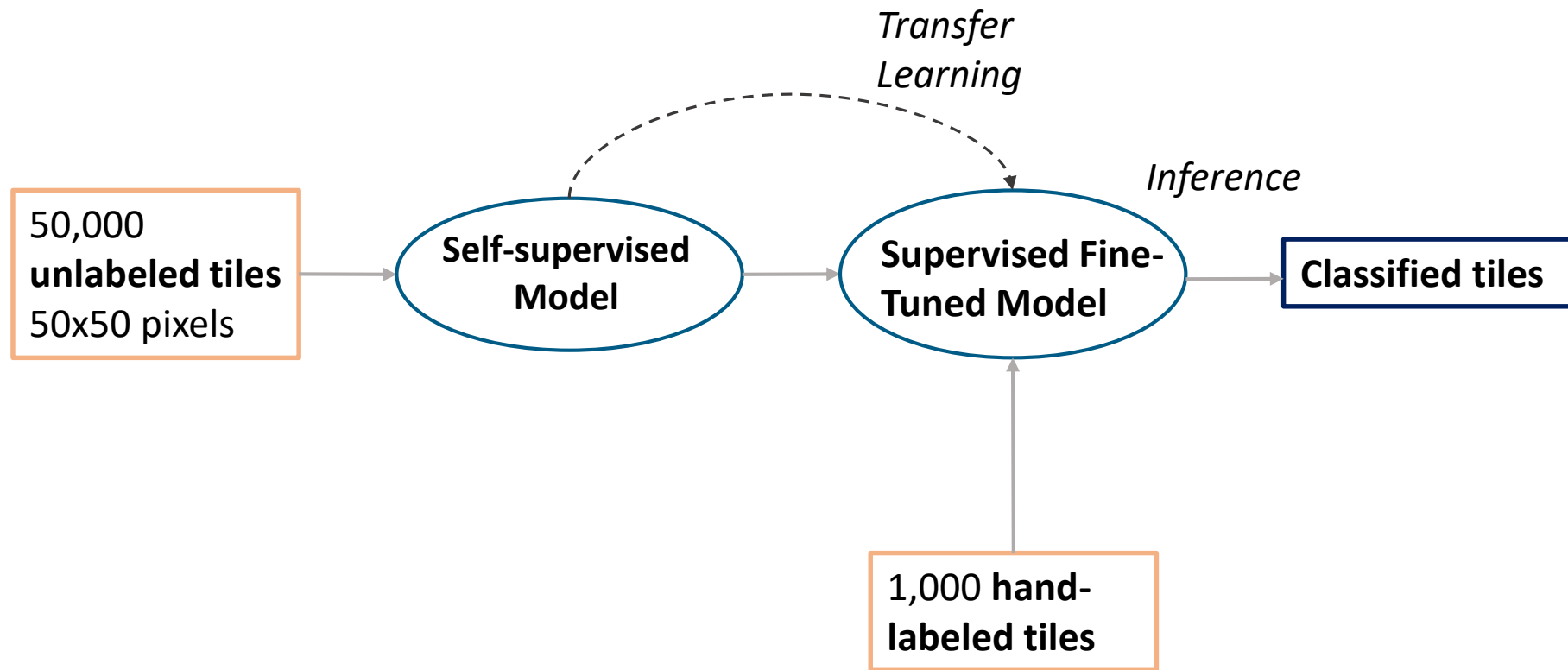
How?

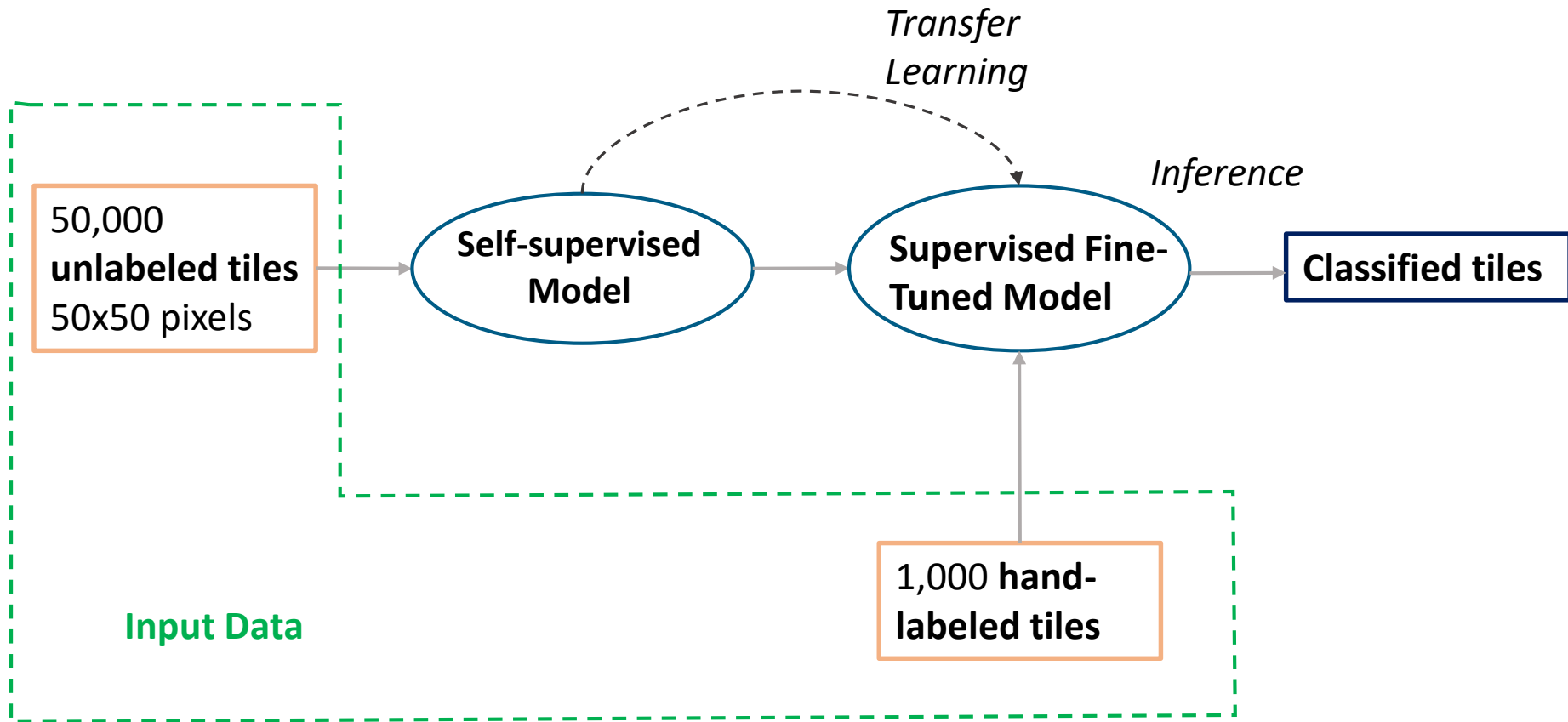
Training a DL pipeline to recognize small tiles of satellite images into one of the following categories:

Bareland **Built-up** **Cropland** **Grassland** **Shrubland** **Tree** **Water**

10 meters resolution images owned and distributed by ESA - European Space Agency



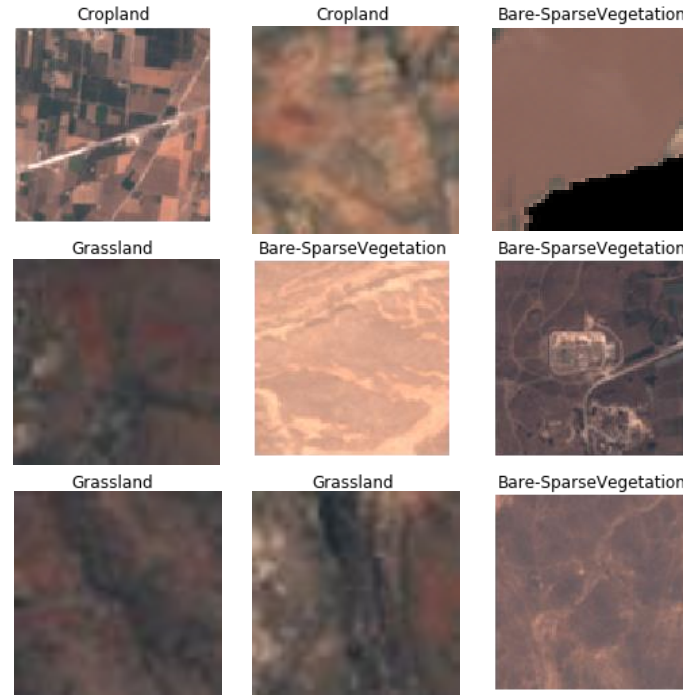
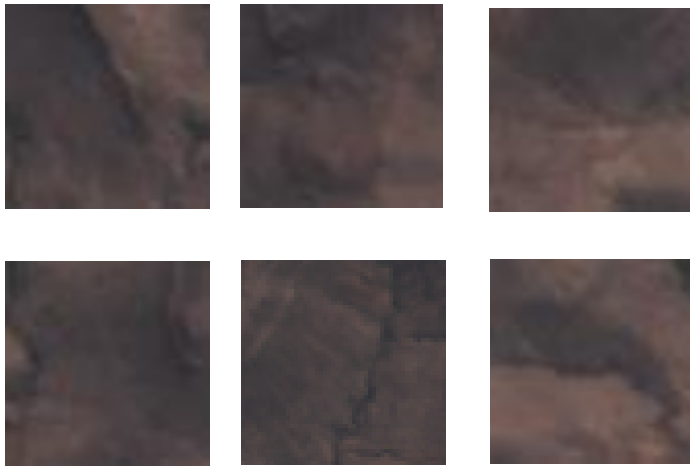




Input Unsupervised Pretrain



Input Supervised Fine-tune



Labels Used

Bareland

Built-up

Cropland

Grassland

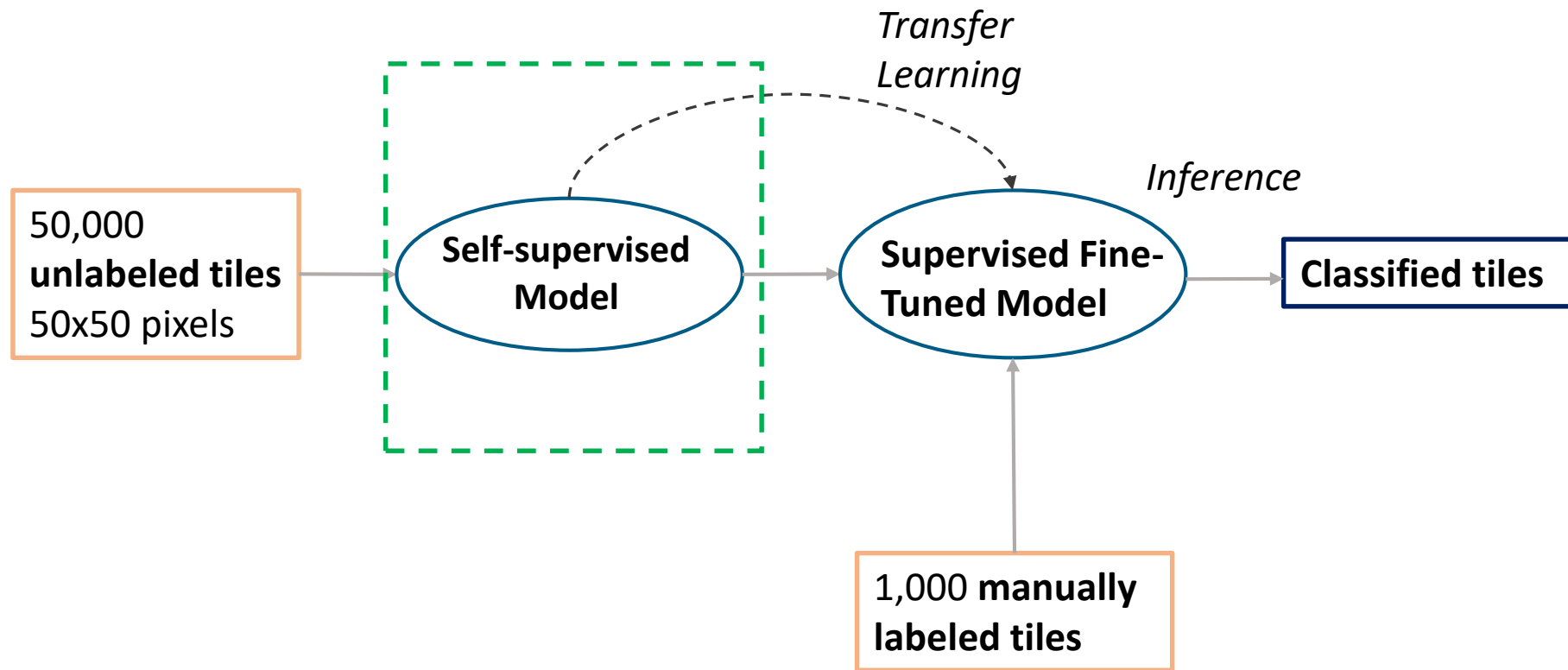
Shrubland

Tree

Water

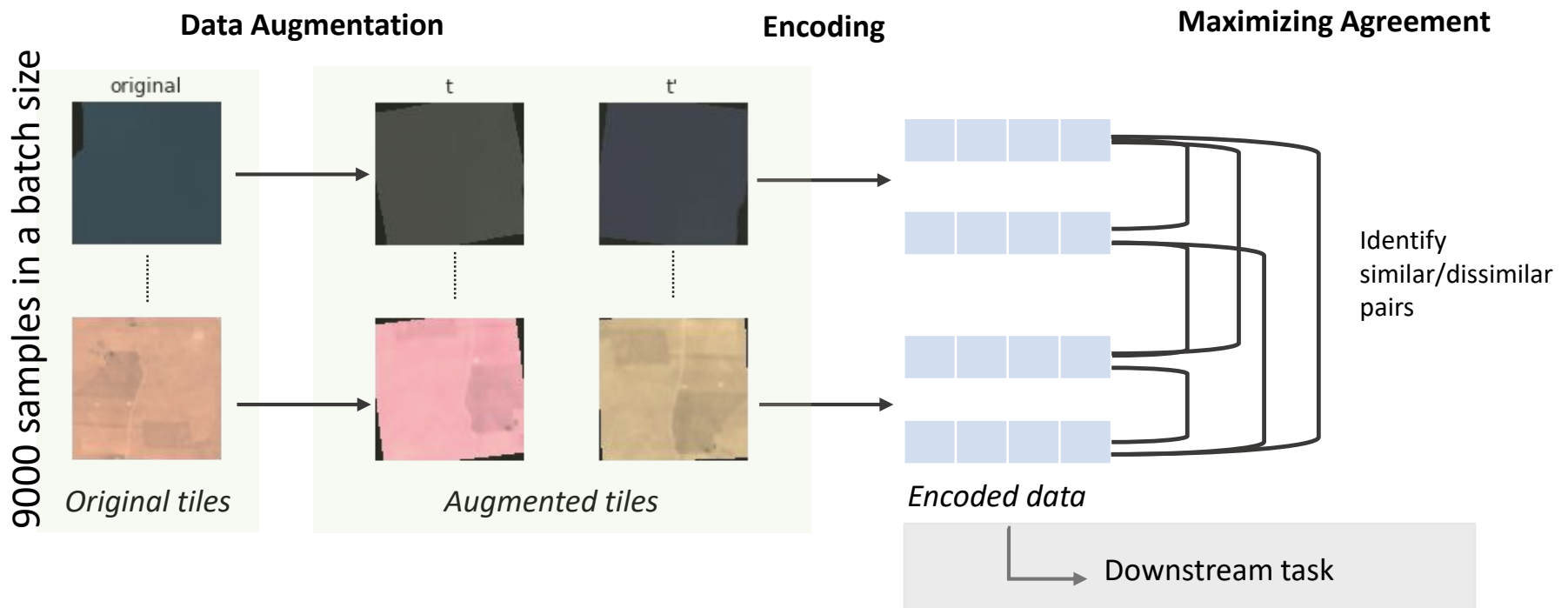
→ Split Sentinel2 acquisition into 50,000 tiles 50x50 pixels

→ once defined 7 labels, a few thousands tiles have been either manually labeled or derived from samples of ESA world Cover



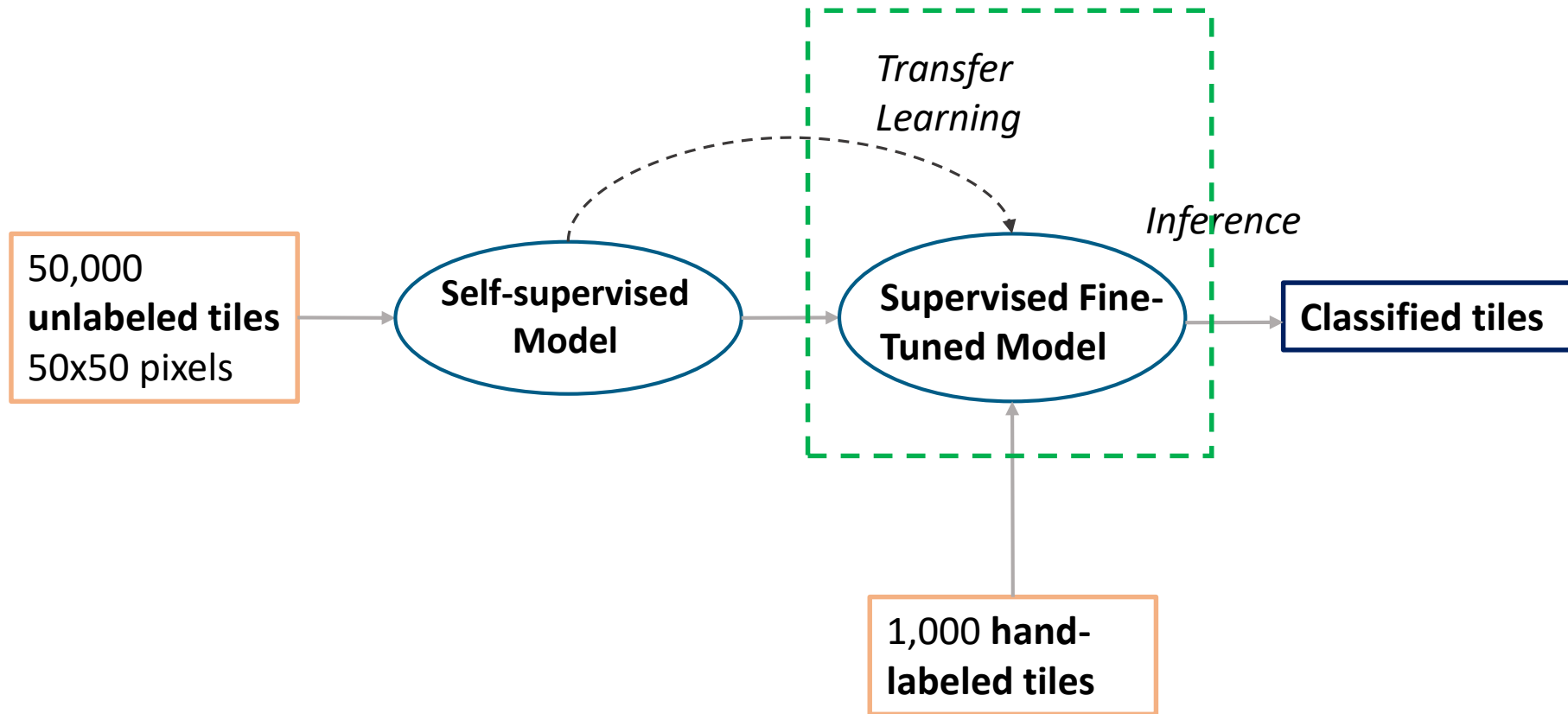
Simclr is a Contrastive Learning technique aiming to learn general features from an unlabeled dataset by teaching a model which data points are similar and which are not.

3 steps: Data Augmentation, Encoding, and Maximizing Agreement.



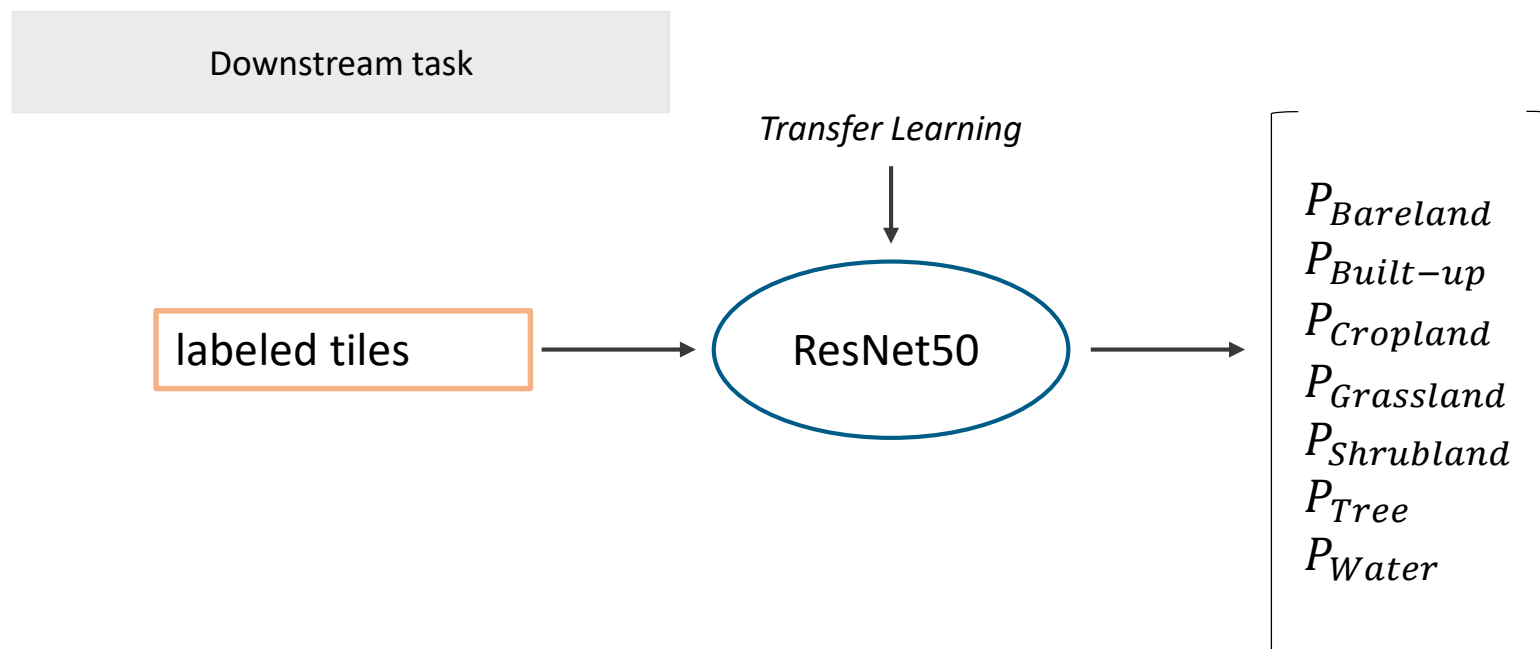
➔ This technique allows to avoid the use of very large quantities of labeled data

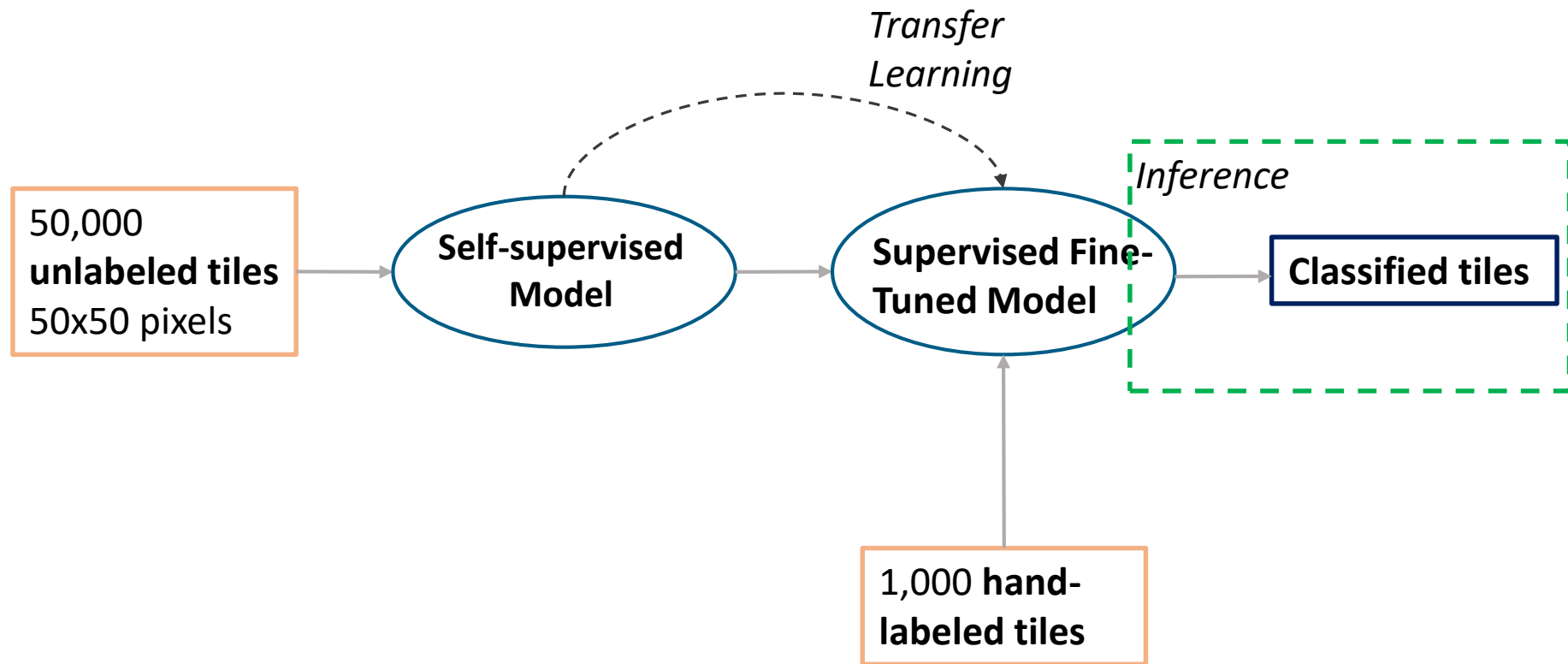
Supervised Fine-Tuned Model



Supervised Fine-Tuned Model

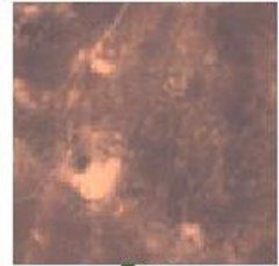
The trained Unsupervised model can be fine-tuned for a specific task → **classification**.
Once the weights from the pretrain are computed, a Model is trained to classify 50x50 pixels tiles into one of the labels.





Sentinel2 tiles classification, a few examples:

Bare-SparseVegetation
Bare-SparseVegetation



Tree
Tree

Grassland
Grassland



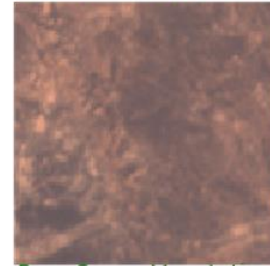
Cropland
Cropland

Cropland
Cropland



Grassland
Grassland

Bare-SparseVegetation
Bare-SparseVegetation



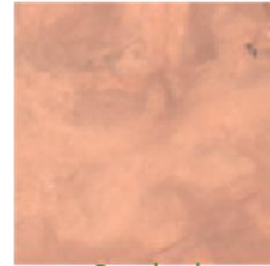
Bare-SparseVegetation
Bare-SparseVegetation

Grassland
Cropland



Cropland
Cropland

Bare-SparseVegetation
Bare-SparseVegetation



Grassland
Grassland



Bare-SparseVegetation
Bare-SparseVegetation



Cropland
Cropland



Water
Water



Cropland
Cropland



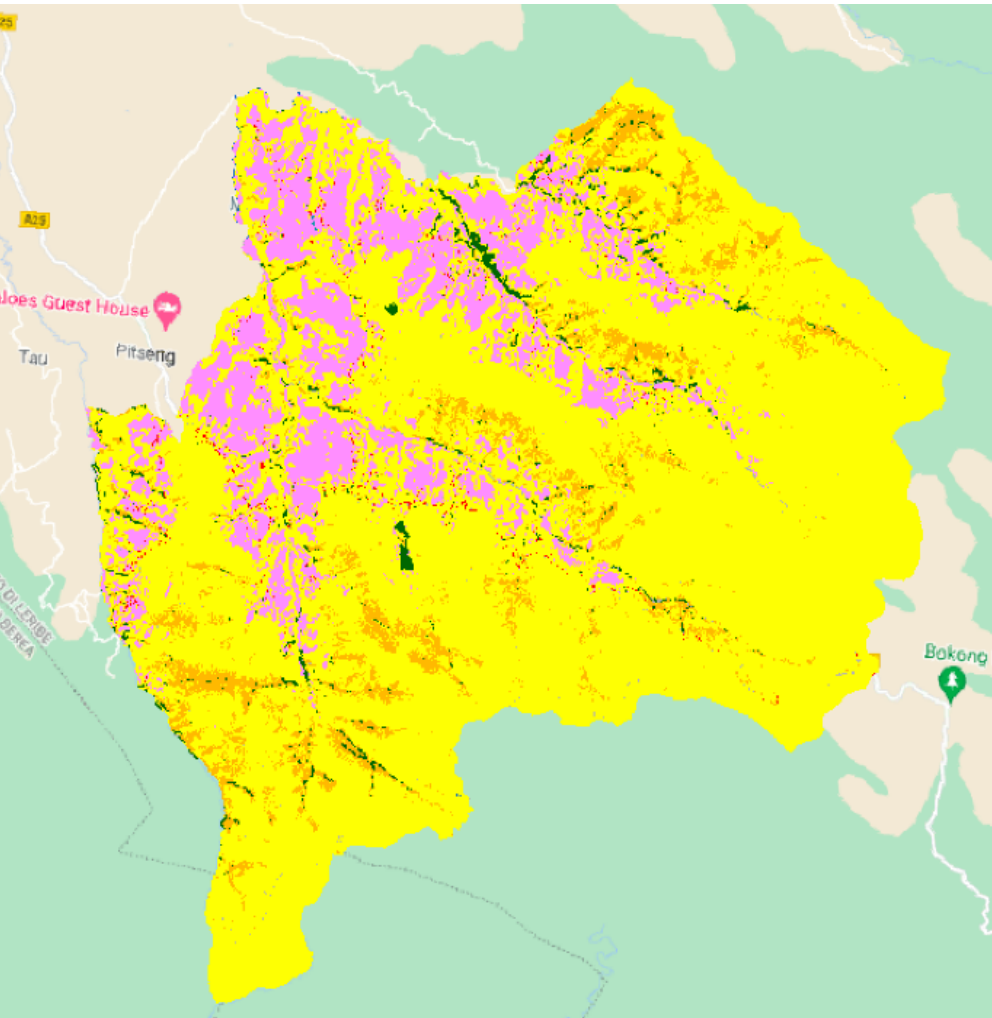
Grassland
Grassland



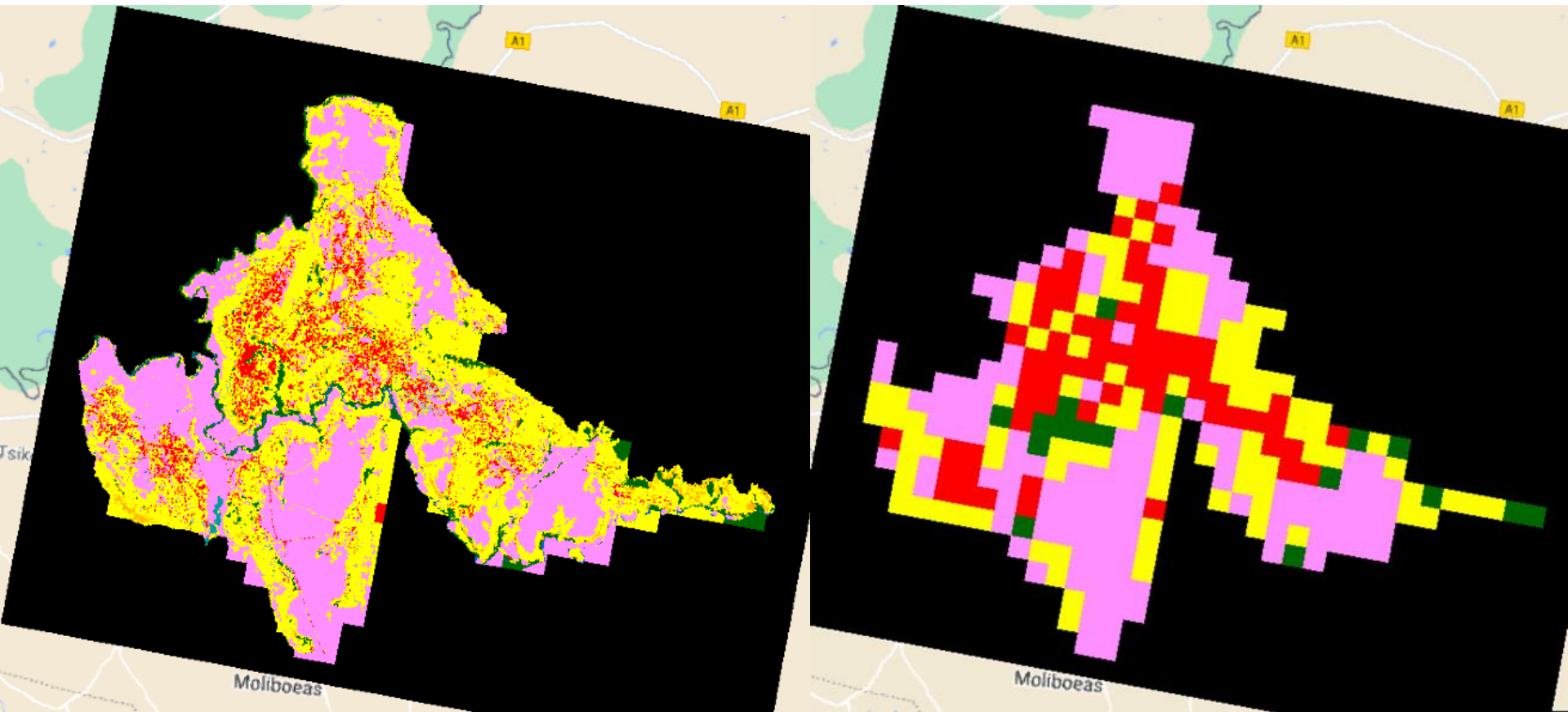
Grassland
Grassland



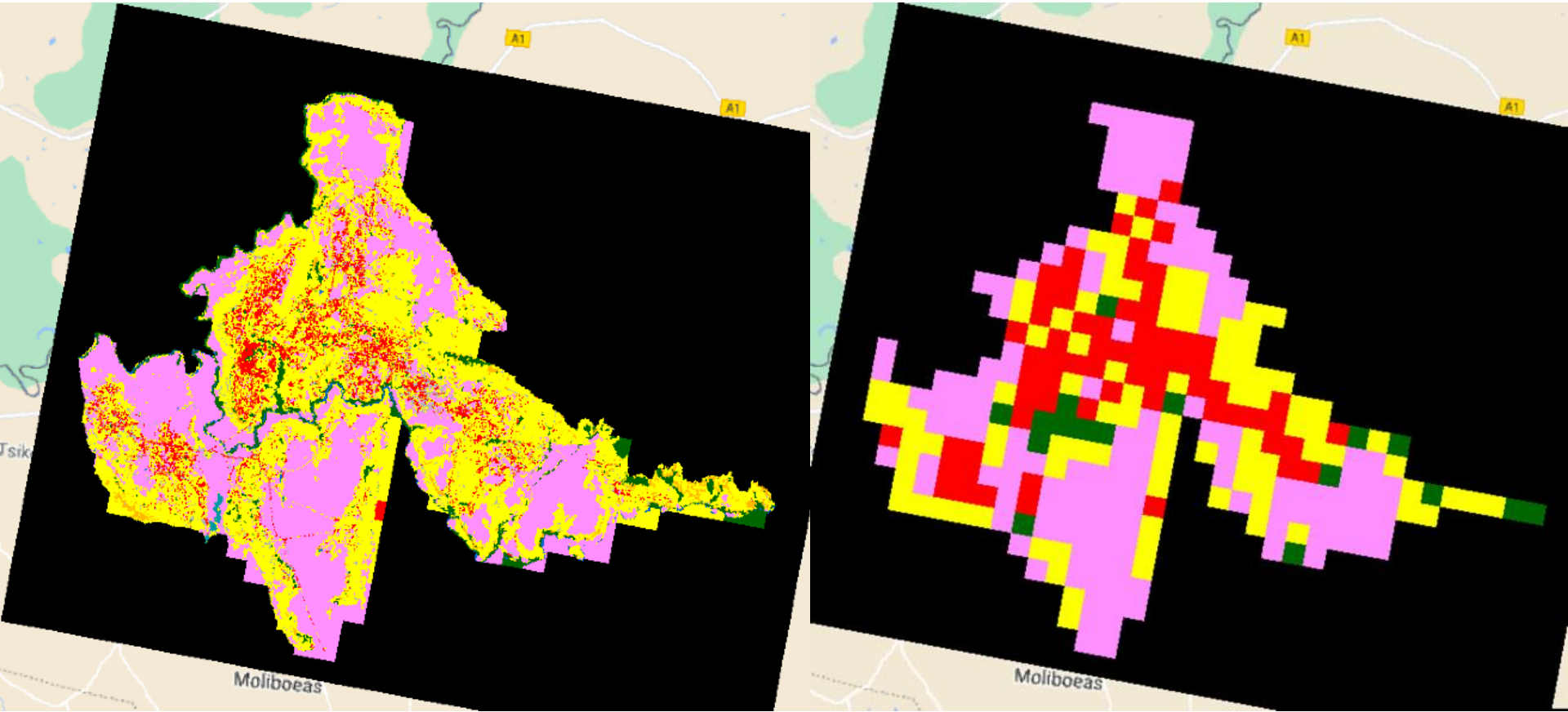
Bolahla, *ESA World Cover* – our R&D work



Leribe, *ESA World Cover* – our R&D work



Leribe, *ESA World Cover* – our R&D work



→ [12,48,24,10,99....], [.20,.35,.10,.30,....]

→ [15,50,20,9,100,....], [.21,.37,.10,.30,....]

Coherences by district and class

District	Coherence with ESA	Coherence with OCS	Coherence with DW
Berea	0,99457	0,92289	0,56838
Butha-Buthe	0,99976	0,95944	0,74237
Leribe	0,99785	0,93260	0,63217
Mafeteng	0,91189	0,98585	0,29727
Maseru	0,99519	0,96304	0,71376
Mohale's Hoek	0,99229	0,98198	0,54063
Mokhotlong	0,99992	0,97339	0,94837
Quacha'sNek	0,99864	0,99389	0,97009
Quthing	0,99647	0,95886	0,64854
Thaba-Tseka	0,99956	0,96667	0,90062

Class	Coherence with ESA	Coherence with OCS	Coherence with DW
Built-up	0,97518	0,90572	0,94782
Cropland	0,97707	0,94429	0,47858
Tree	0,79038	0,81077	0,64517
Shrubland	0,50909	0,40835	0,76543
Grassland	0,99620	0,99799	0,95165
Water	0,80179	0,80817	0,48538

PROs

CONs

the entire pipeline
made in a few
weeks

Possibility to
finetune

output weights can
be easily transferred
for inference

Need GPU
infrastructure



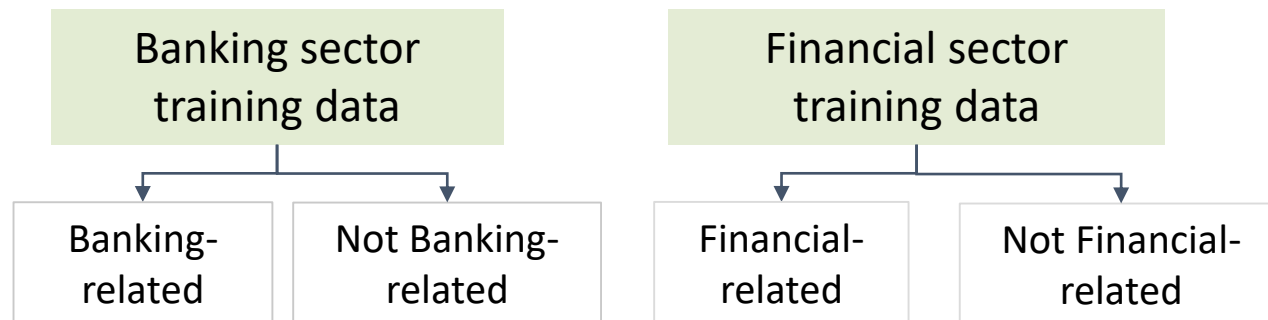
Case scenario 2 – Natural Language Processing

*Gain insight into **national and international opinion** and **popularity** with respect to the **economic sector***

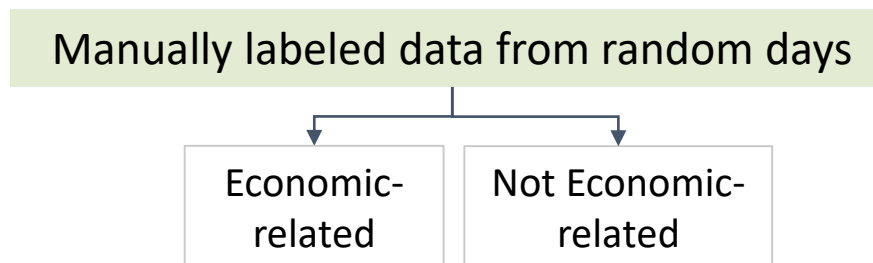
How?

Training a DL pipeline to identify economic/non-economic-related article contents

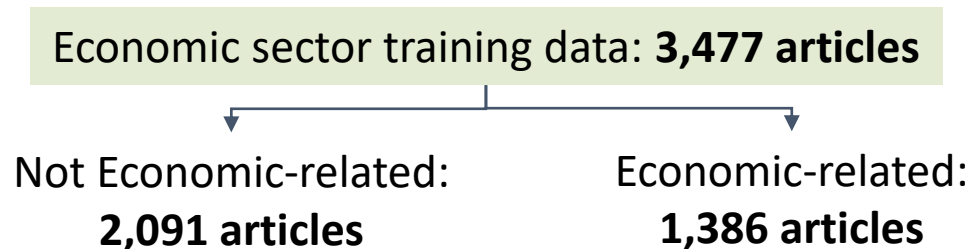
Step 1: Re-label in-house training data to the problem at hand



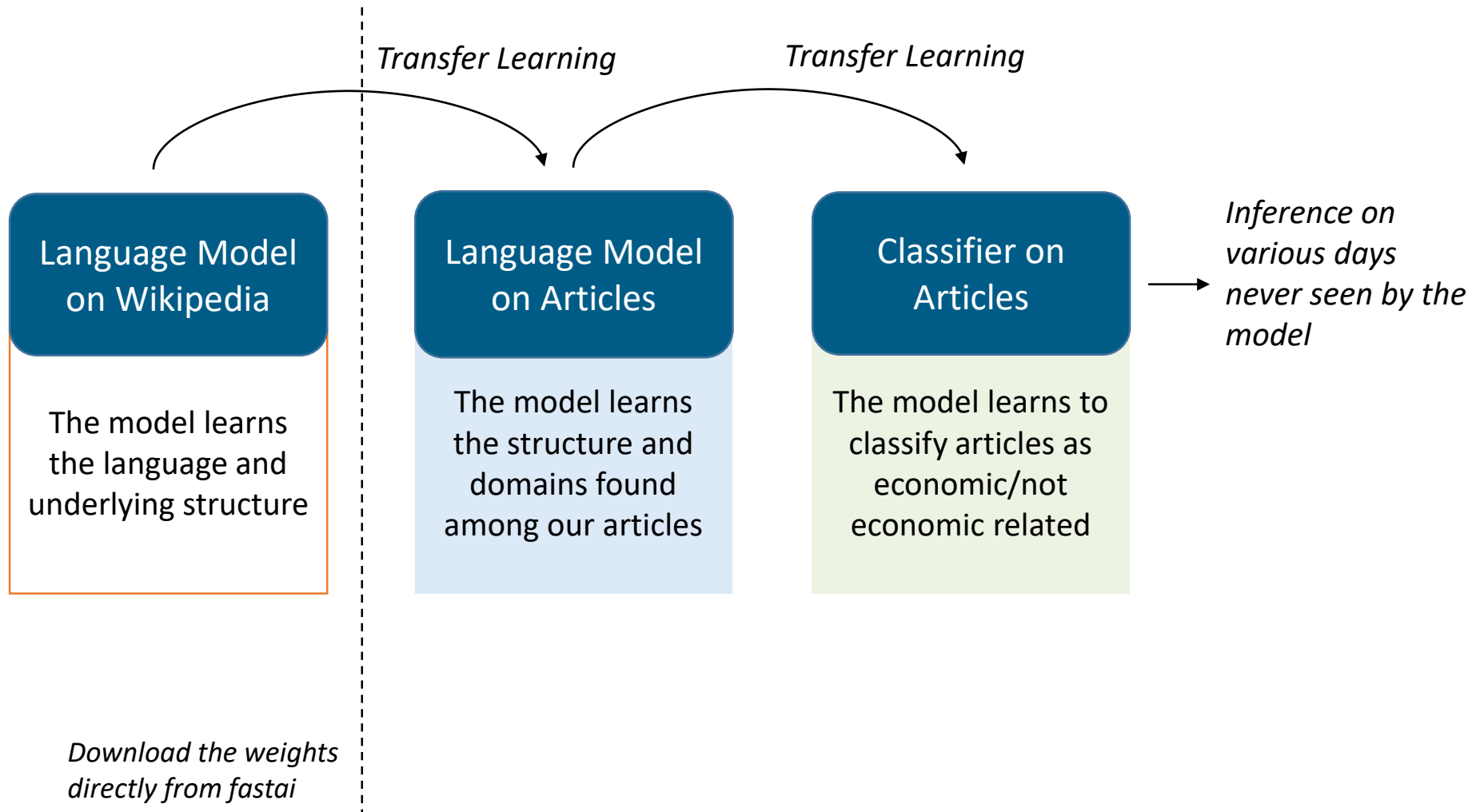
Step 2: Add manually labeled data



Final Training Data



ULMFiT approach (Universal Language Model Fine-Tuning), embedded in fast.ai



Reviewed all the articles from a randomly picked day...

		<i>Actual</i>	
		6/7/22	no
<i>Predicted</i>	yes	253	43
	no	59	10,076

Precision: 85,47 %

Recall: 81,09 %

... And from other days during that week

Date	Total Articles	Predicted Yes	FP	Precision
7/7/22	10,337	322	29	90.99%
8/7/22	9,795	250	7	97.20%
9/7/22	6,856	141	13	90.78%
10/7/22	6,683	110	13	88.18%
11/7/22	5,725	158	9	94.30%

- Monetary operations (interest rates, inflation)
- Fiscal operations (tax, spending, debt, deficit)
- Labour market (employment, wages)
- Consumers'/investors' confidence
- Trade
- Housing market
- ... any event that has an effect on a country's economic health

Predicted - yes

- **"Noisy" topics** that are linked to the economy:
 - tourism sector (aviation industry)
 - Performance of single companies
 - Increases in food prices (food inflation), and fuel affecting the cost of living

Actual - yes

Actual - no

- Articles that are **too short**
- Articles that contain a lot of **advertisements**

Predicted - no

- Stranger Things Season 4
- K-Pop
- Floods
- War in Ukraine
- Covid-19 & variants, monkey pox
- ... Basically everything else!

Correctly identified as non-economic related

Debt funds have a **portfolio of securities** and hence help you diversify when compared to a single instrument like a **fixed deposit, corporate bond** or a **non-convertible debenture**. <https://t.co/bCdQyldSfr> All you need to know about investing in **debt mutual funds**. All you need to know about **investing in debt mutual funds**. By , ET BureauLast Updated: Jul 06, 2022, 09:34 AM ISTFont Size. Debt funds have a **portfolio of securities** and hence help you diversify when compared to a single instrument like a **fixed deposit, corporate bond** or a **non-convertible debenture**. A sharp rise in **bond yields, liquidity**, indexation benefits and ability to earn more than traditional deposits are driving investors to debt funds. What is a debt mutual fund scheme? A debt fund mutual fund scheme is one that invests in fixed income yielding instruments. Depending on the type of debt scheme, the instruments could be T bills, bank CDs, **corporate bonds, government securities** or a combination of these instruments. Some of the debt fund categories available to investors are overnight, liquid, ultra short term, short term, medium duration, target maturity funds, credit risk, dynamic bond and government securities fund. What is the advantage of a debt scheme over a traditional fixed income instruments? Debt funds have a portfolio of securities and hence help you diversify when compared to a single instrument like a fixed deposit, corporate bond or a non-convertible debenture. A portfolio of a debt scheme helps reduce risk as it has a number of papers. A debt mutual fund is liquid as investors can buy or sell it with the fund house on any working day. If there is a need to withdraw money, a debt mutual fund can be broken into units of `1 and investors can withdraw only the amount required. As compared to this, in a **small saving product** or a fixed deposit, you would need to break the entire deposit. What are the **tax benefits** for an **investor** in a debt fund? Tax benefits are one big reason for investors to prefer debt funds. Firstly, there is no tax deduction at source (TDS) in debt mutual funds and if held for three years and above, one can avail indexation benefit and minimise their tax outflow. Fixed deposits do not offer indexation benefits. What are the risks investors need to keep in mind while investing in a debt fund? Investors need to know about **interest rate risk** and **credit risk**. In a rising rate scenario, bond prices fall and vice versa, which could give mark-to-market losses in the short term. A credit risk is the risk of default on a debt security that may arise from a borrower failing to make required payments. If the scheme holds lower-rated securities and the company does not repay principal or interest on the due date, the fund's NAV could drop, leading to a loss to investors.

PROs

CONs

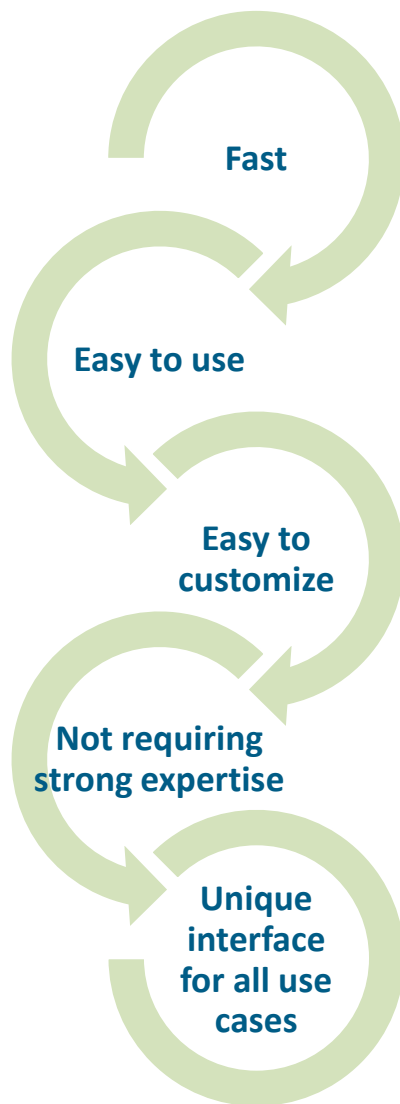
The pipeline was built **very fast**

Architecture can easily be used for **other domains/topics** or adapted to **multi-class problems**

Can be **costly** for use cases that need to display results in a timely manner

Need **GPU** infrastructure

To wrap up, fast.ai is...



Q&A