Using Neural Networks to Classify 3D Scans of Museum Artifacts

Capstone Project in Machine Learning, 2021

Alvaro Ortiz Troncoso

EPFL

https://www.epfl.ch/

3D-digitalization before 2021

Automated 3D scanning at Museum für Naturkunde Berlin, using the CultLab3D system developed by Fraunhofer institute (Photo: Carola Radke, MfN)

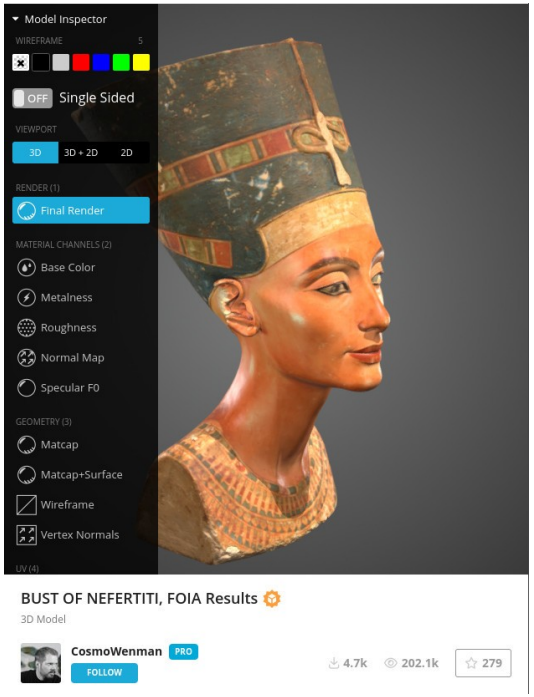
3D-digitalization after 2021

Automated 3D scanning at Museum für Naturkunde Berlin, using the collection digitization system developed by Picturae (Photo: Maren Demant, ifs)

Given the current state of digitization of cultural artifacts,
what results can be expected?

Until 2021

Digitizing an artifact can take several hours (including post-processing), therefore digitization reserved for *the most prestigious items* in the collection



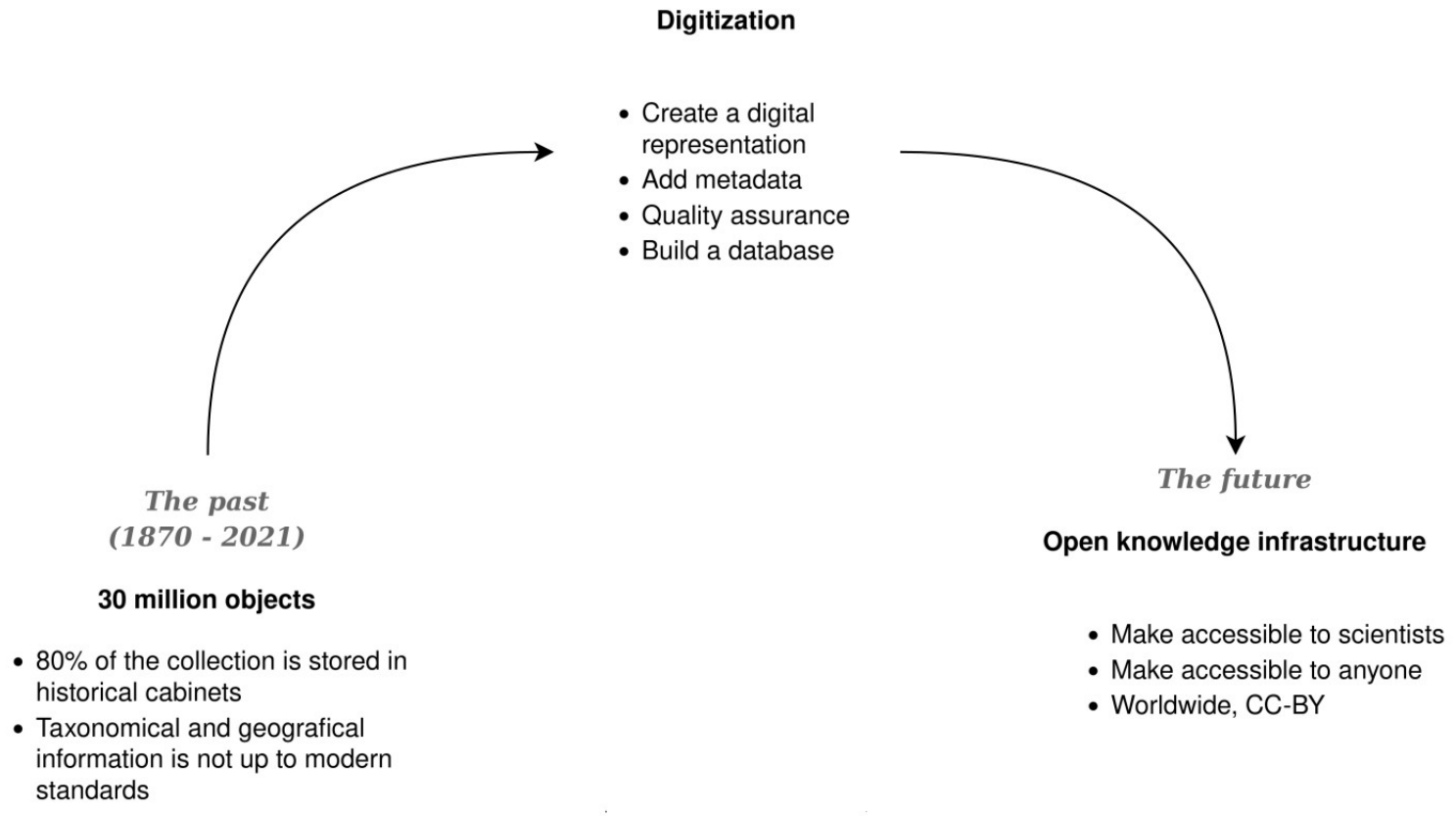
3D-Model by Trigon Art, Berlin 2008 CC BY-SA,
Staatliche Museen zu Berlin – Preußischer Kulturbesitz

After 2021

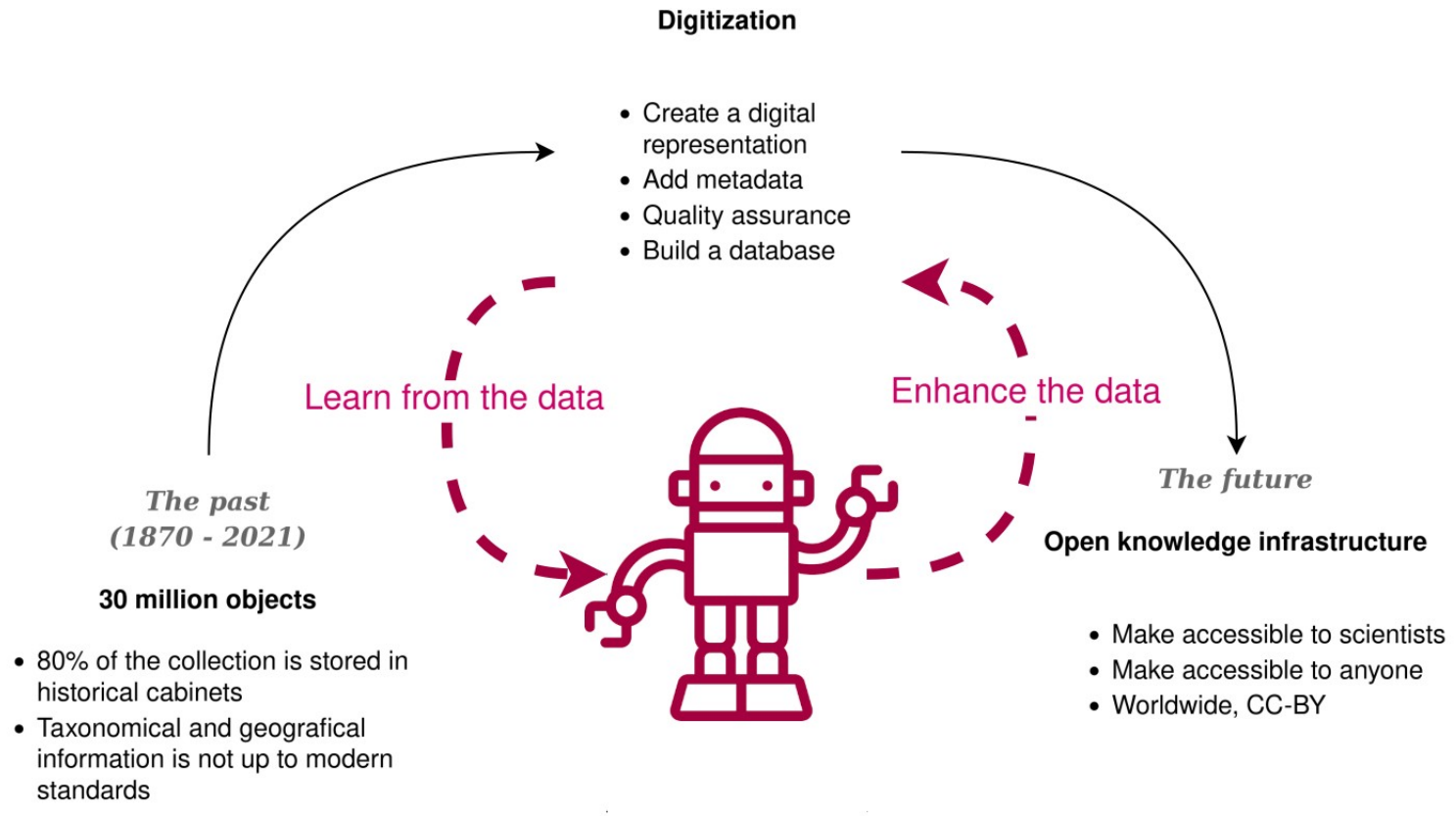
Cost-efficient and timely mass-digitization of *all*
the artifacts in a collection

(Image: Michael Scheuerl / MfN)

What problem could be solved by mass digitization?



What problem could be solved by mass digitization **+ Machine Learning**?



Enhance incomplete metadata by learning from complete metadata.

Data Sources
Digitization projects in Archaeology

University of Virginia, Charlottesville, USA

Musée Art & Histoire, Bruxelles, BE

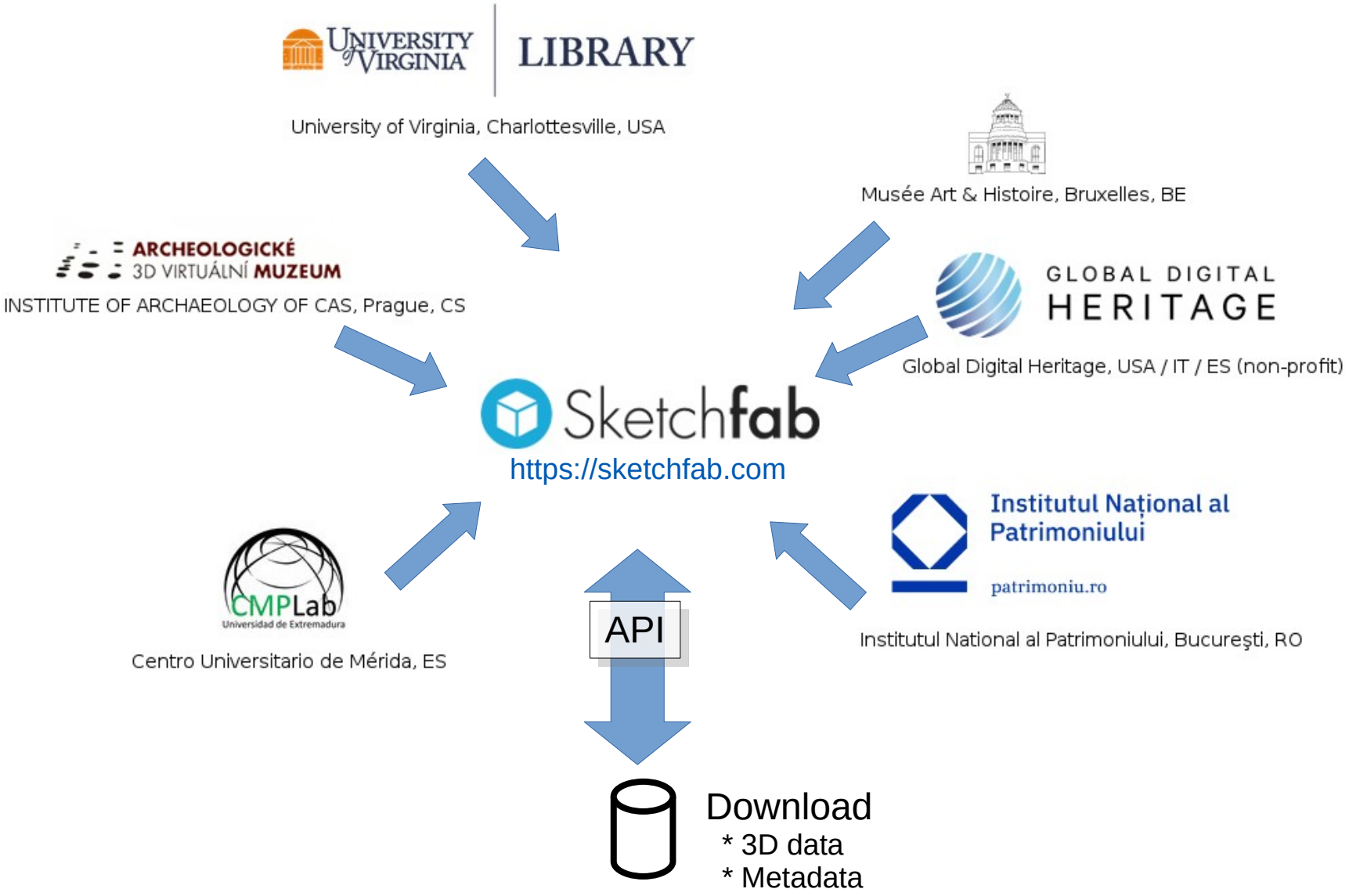
INSTITUTE OF ARCHAEOLOGY OF CAS, Prague, CS

Global Digital Heritage, USA / IT / ES (non-profit)

Centro Universitario de Mérida, ES

Institutul National al Patrimoniului, Bucureşti, RO

Data Sources Aggregator

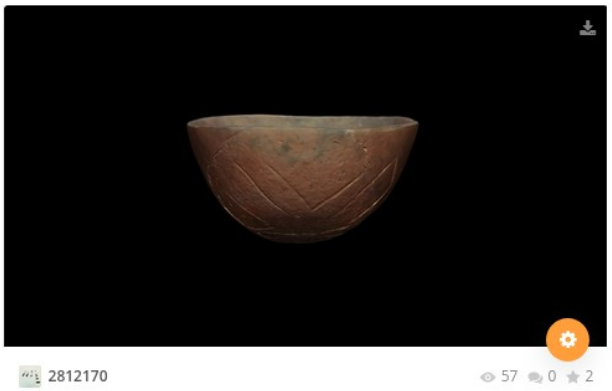


Downloaded data: Metadata

1) the name of the 3D model
2) ideally a textual description of the object, written by a museum curator
3) the API-URI where the 3D-model file can be downloaded
4) a link to the 3D-model's preview page on Sketchfab
5) the name of the museum, institution or organization that provided the digitized object
6) a link to the providers page on Sketchfab
7) the license of the 3D-model
8) the number of vertices in the 3D-model
9) tags added by the provider of the 3D-model describing the object (json)

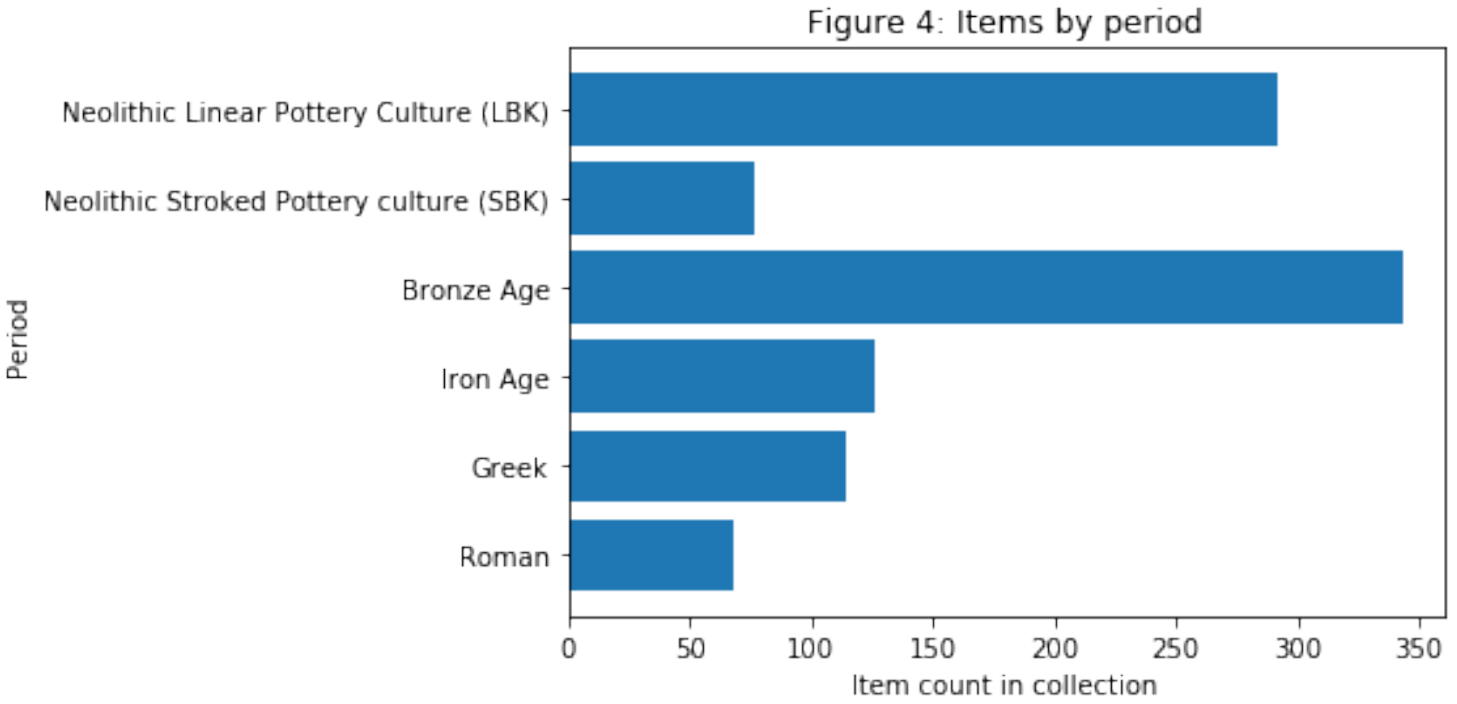
	Data Provider	Object Count
1	Archaeological 3D virtual museum	369
2	The Royal Museums of Art and History	343
3	Global Digital Heritage	180
4	Institutul Național al Patrimoniului	22
5	CMPLab	11
6	UVA3D	10
7	Other (less than 10 models each)	85

Downloaded data: Classes

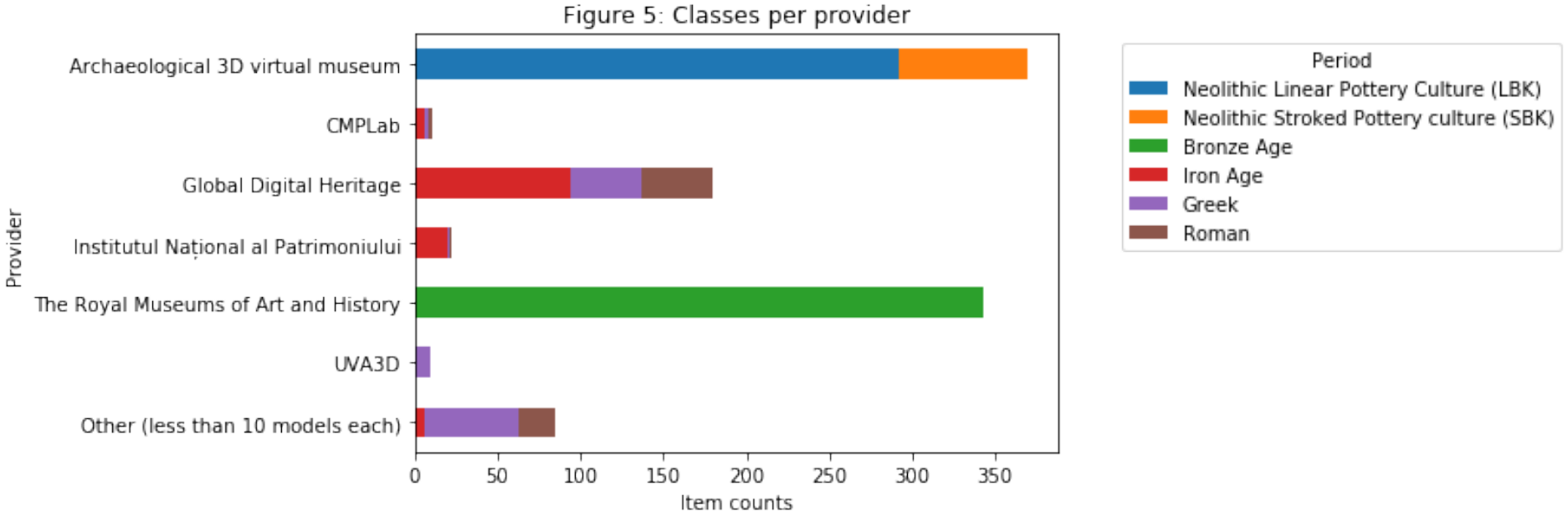


Various types of pots in the data set: Neolithic LBK vessel, Neolithic SBK mortar, Bronze Age bowl on foot, Iron Age -Visigoth- bottle, Greek Jug, Roman cup (screenshot from the Sketchfab website).

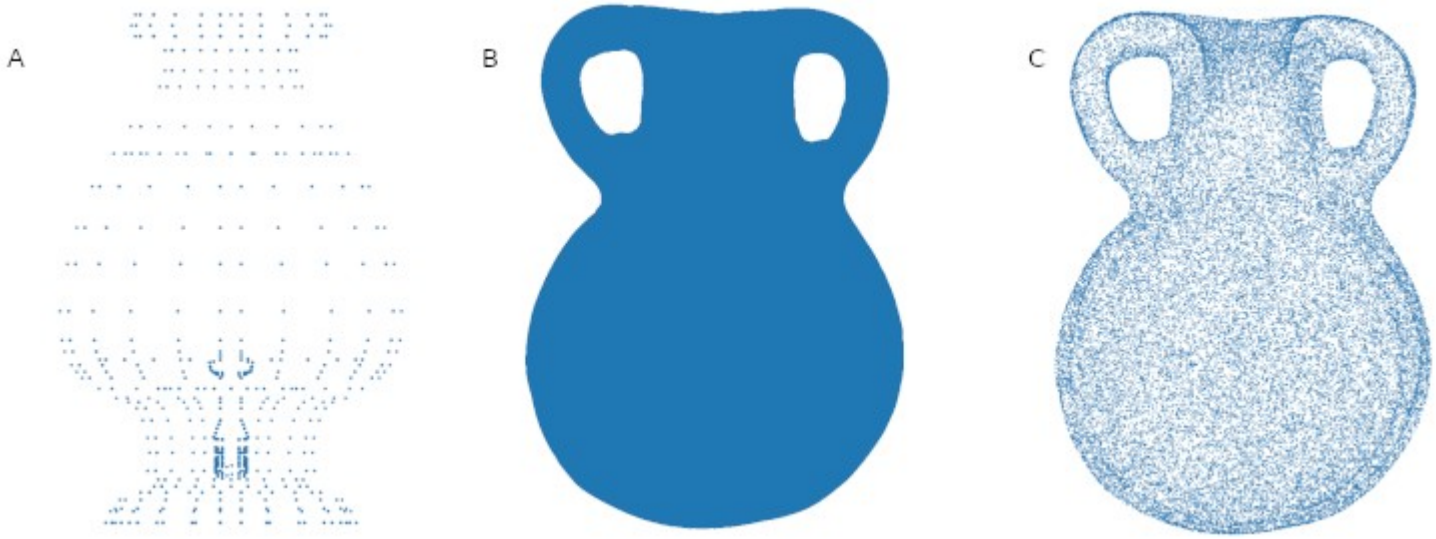
Caveat 1: Data is imbalanced by class



Caveat 2: Data is imbalanced by source

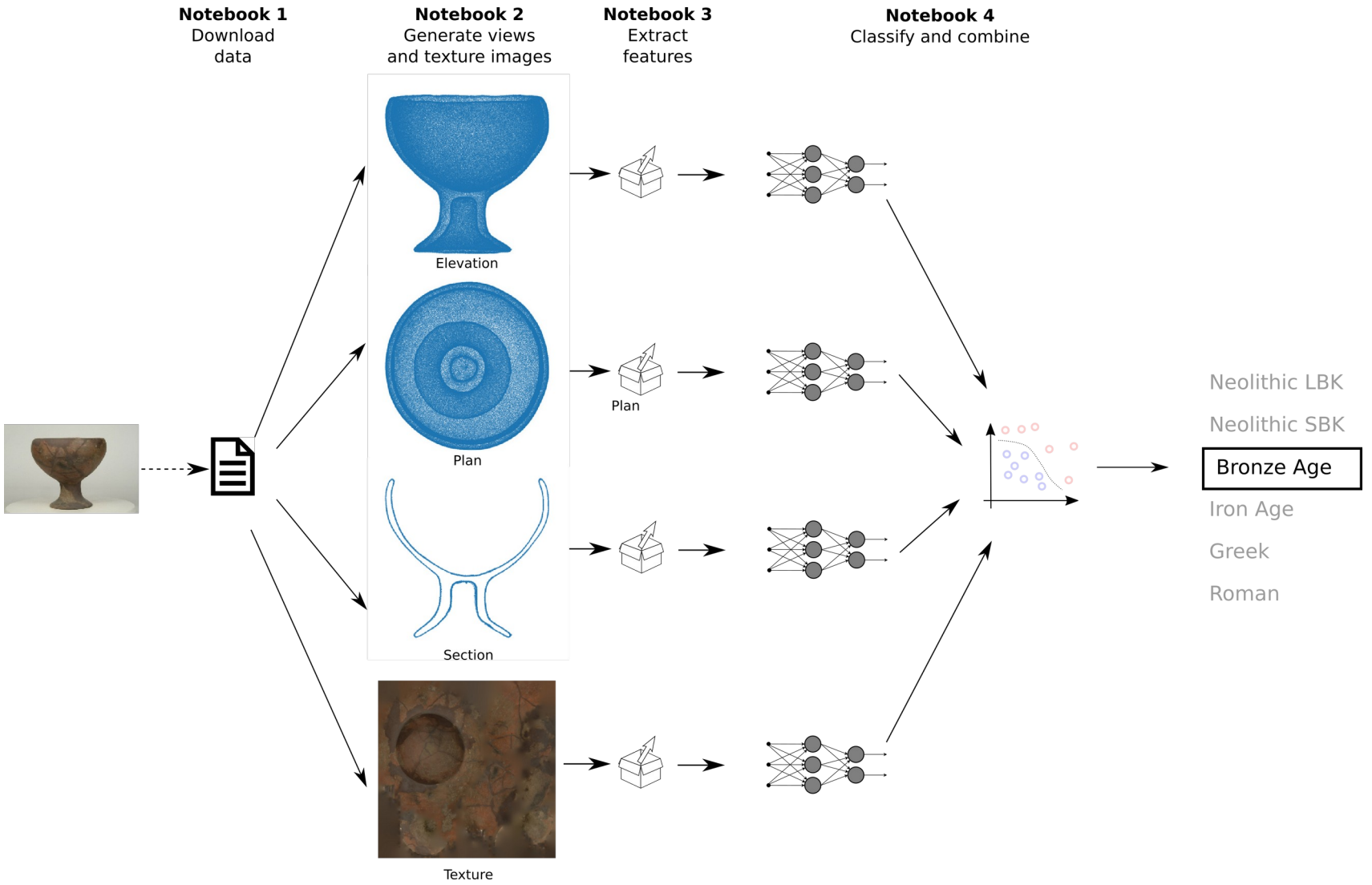


Downloaded 3D models have to be homogenized



The 3D-model with the least vertices (A), with the most vertices (B), the model with the most vertices sub-sampled to 37532 vertices (C)

Machine learning setup



Preprocess 3D-models

A Greek jug (A), its 3D point cloud (B), its texture image (C).

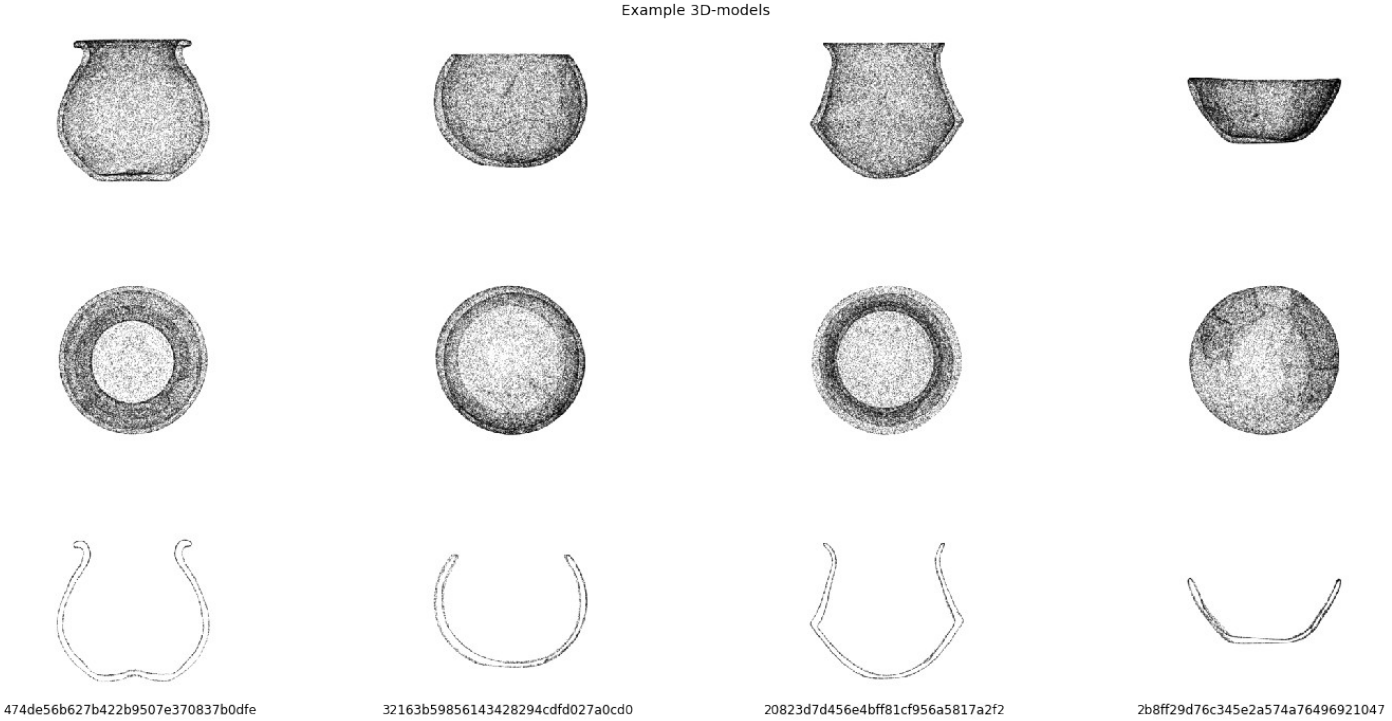
Point clouds

1) sub sample point cloud data so that the number of vertices is the number computed in the EDA in notebook 1, i.e. 37838 vertices
2) rotate the point cloud to generate an elevation (side) view
3) generate plan (top) and cross-section views by rotating the elevation view
4) render the views as images
5) save the pixel data of the flattened images into a npz file

Textures

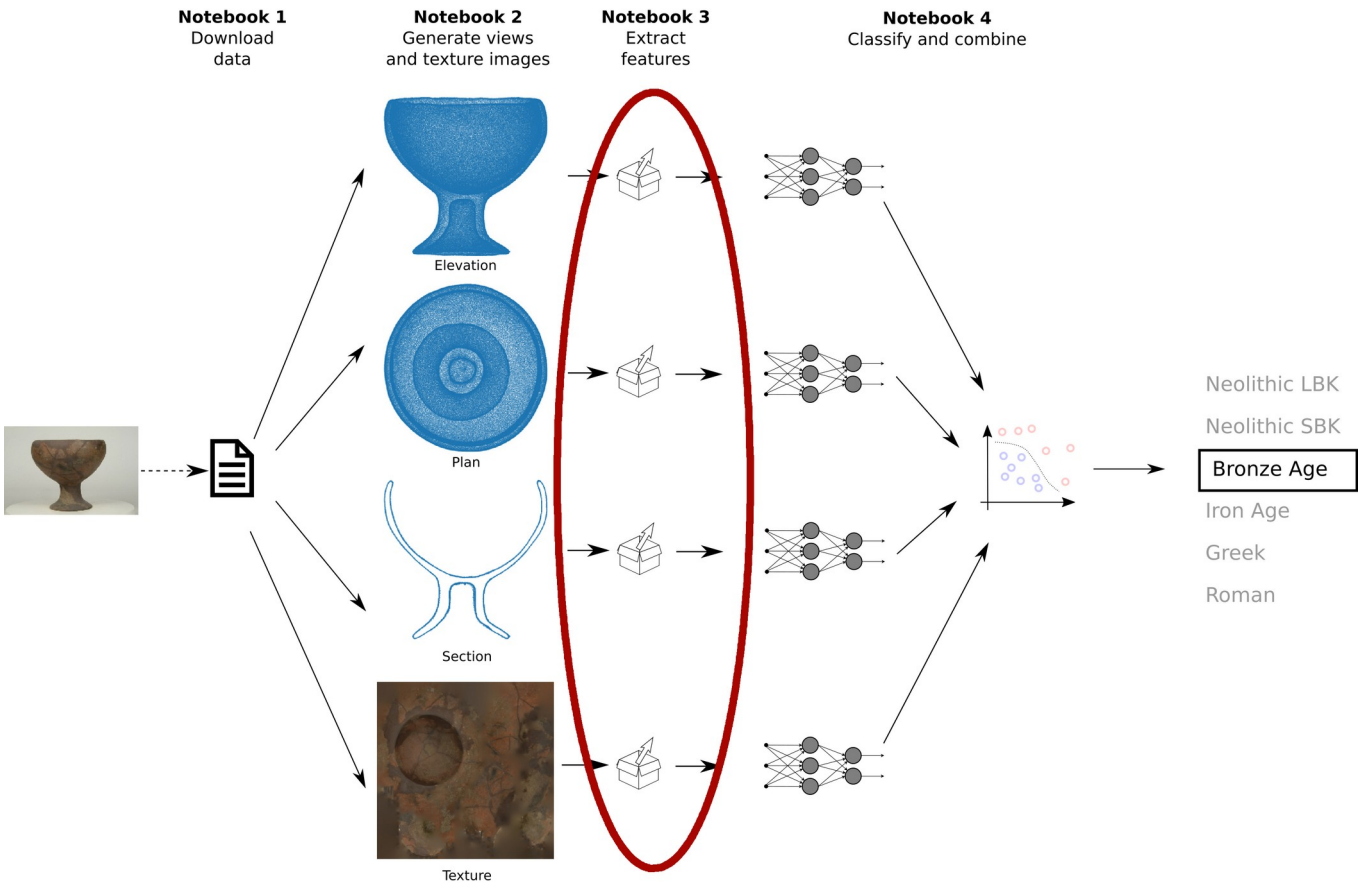
6) resize the texture image to the size expected by the data extraction model, i.e. 299x299 pixels RGB
7) save the textures as flattened images into a npz file

Split data into "train", "validate" and "test" datasets

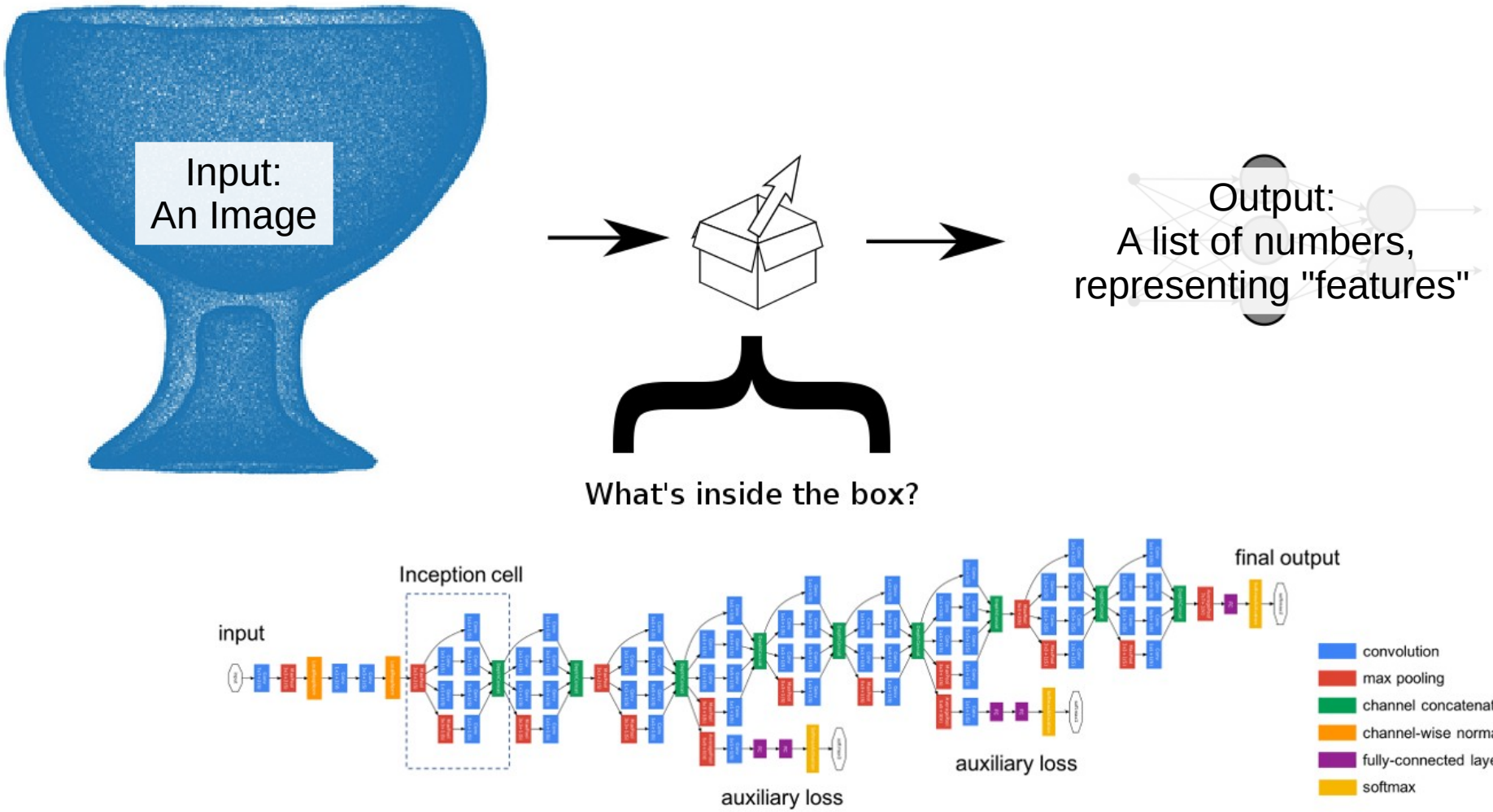


- Training dataset (70%)
- Validation dataset (20%)
- Test dataset (10%)

Extract features by transfer learning



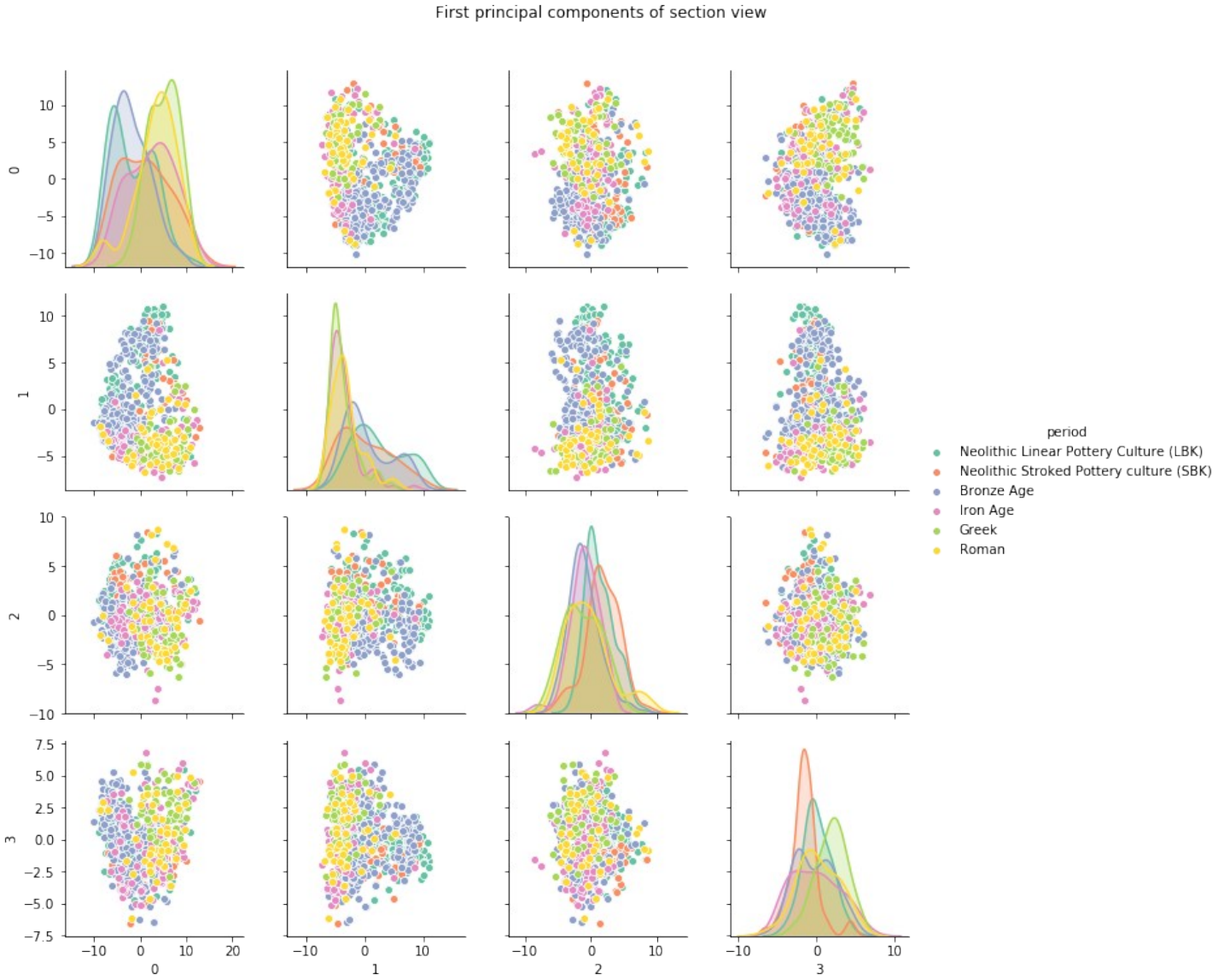
Extract features by transfer learning



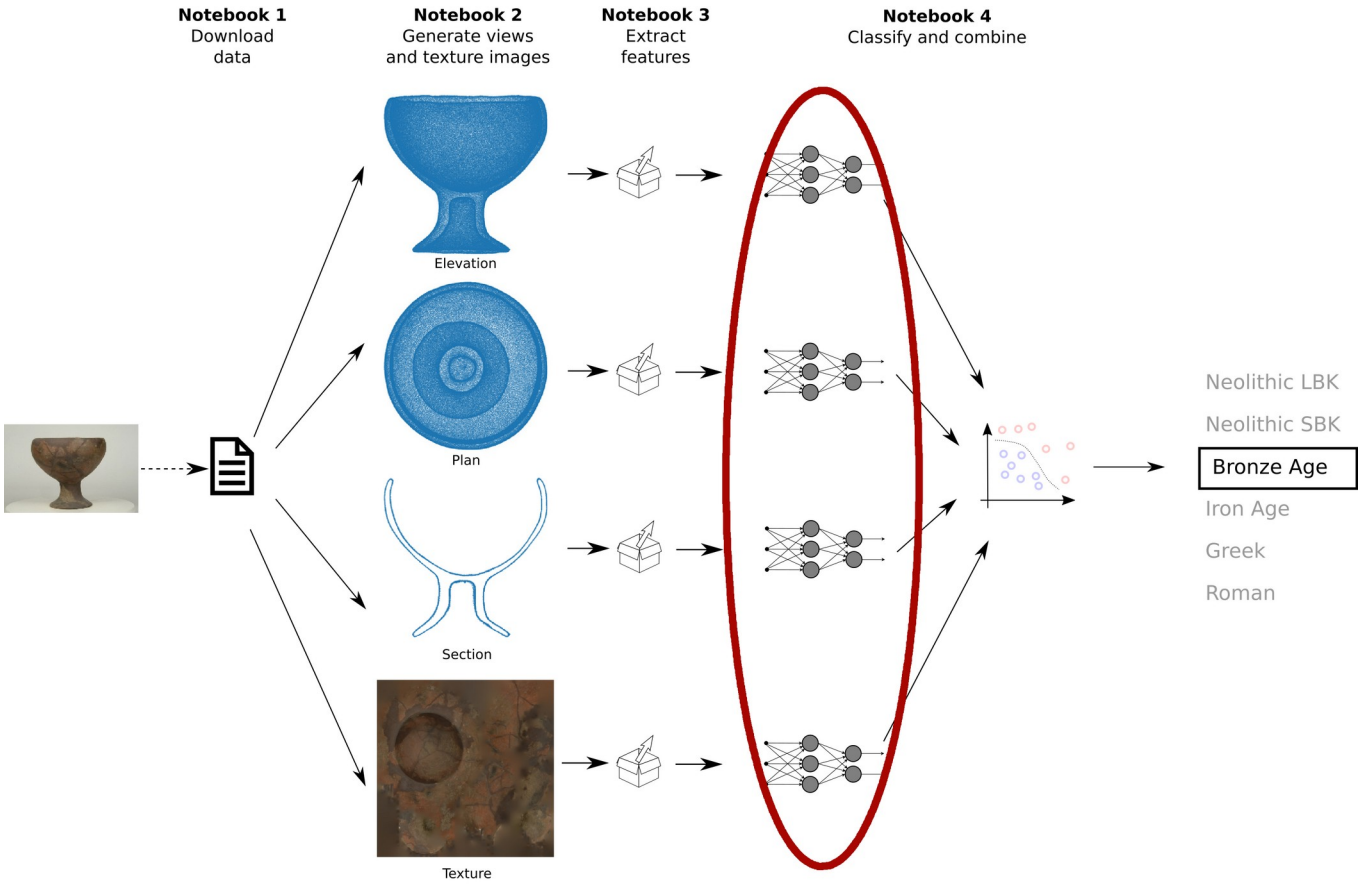
Inception v3 model, trained on "ImageNet" dataset (14.197.122 images) on Google Cloud

Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jonathon Shlens, Zbigniew Wojna: "Rethinking the Inception Architecture for Computer Vision", 2015.

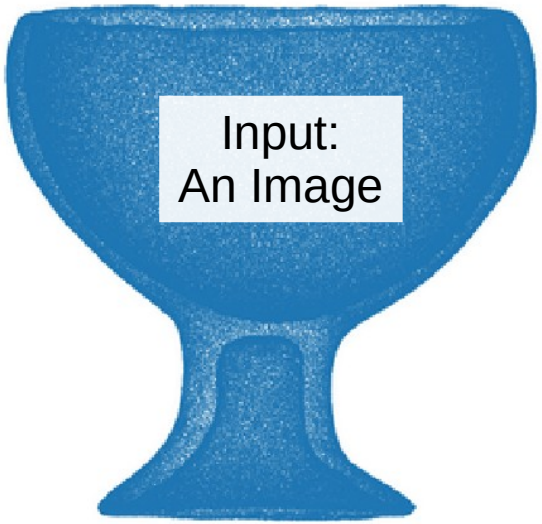
Principal component analysis

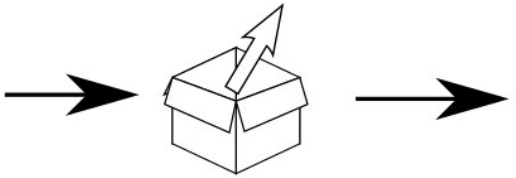


Classify

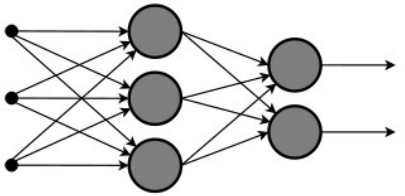

Notebook 1
Download
data
Notebook 2
Generate views
and texture images
Notebook 3
Extract
features
Notebook 4
Classify and combine
Elevation
Plan
Section
Texture
Neolithic LBK
Neolithic SBK
Bronze Age
Iron Age
Greek
Roman

Classify





Intermediate result:
A list of numbers,
representing "features"

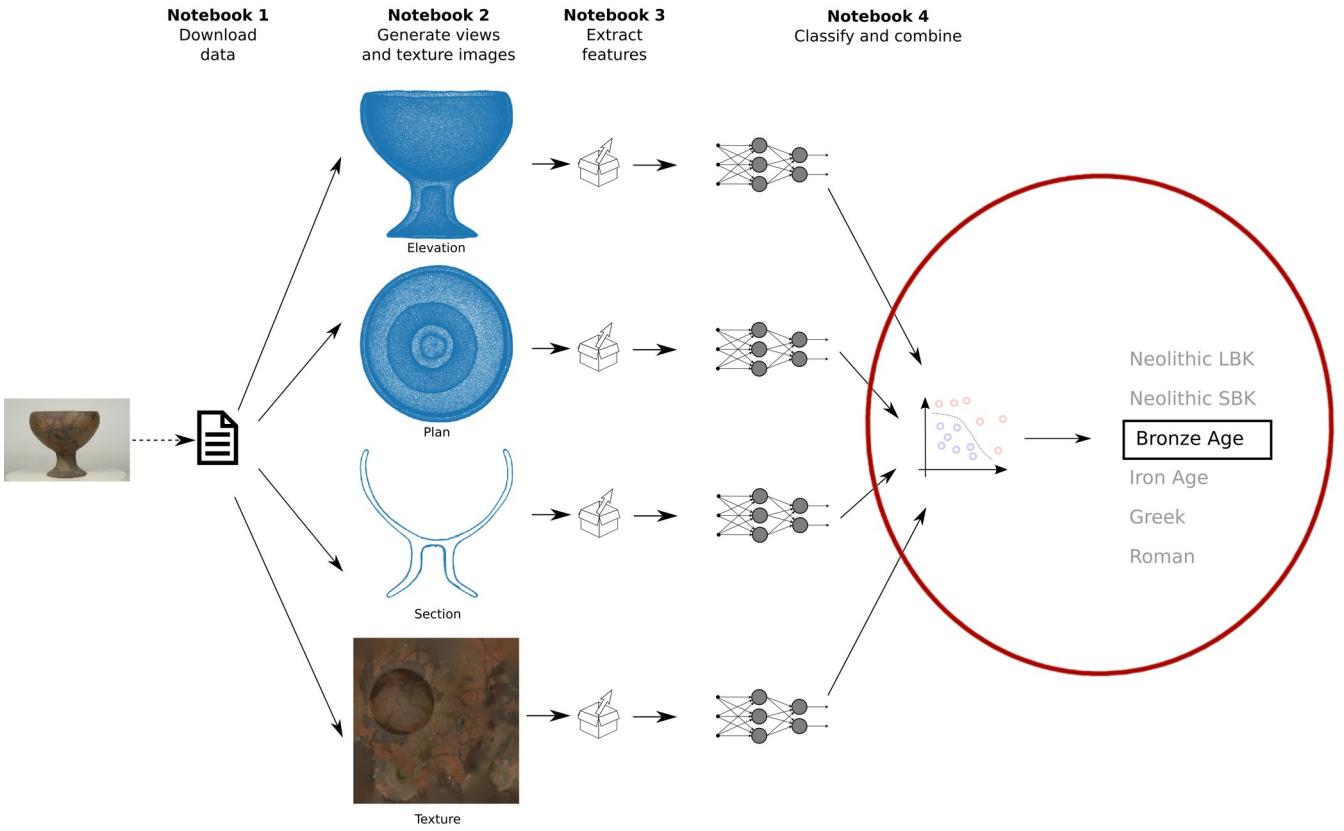


Output:
Probability that the artifact
belongs to each of
the 6 classes

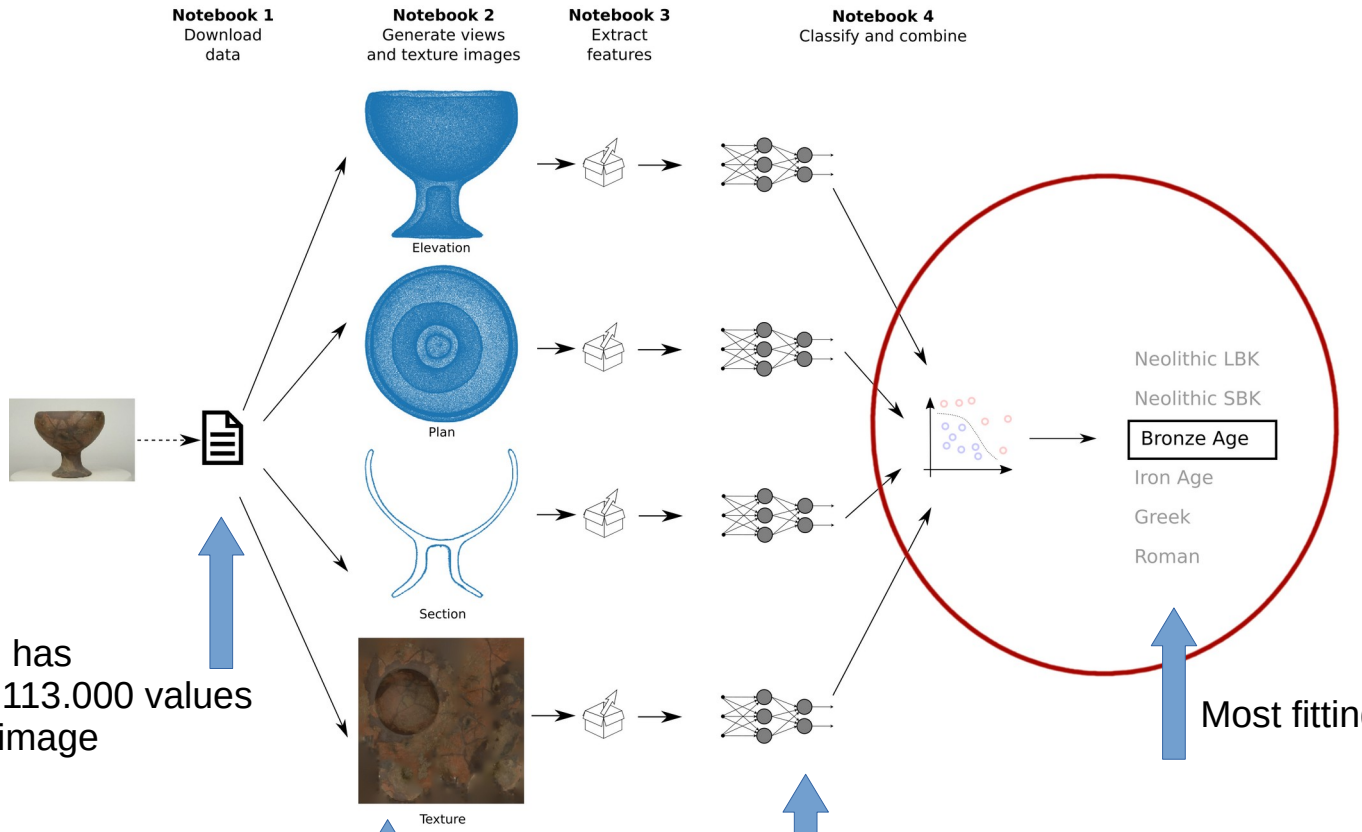
Example:
For the first artifact, for the "elevation view"

Neolithic LBK	50%
Neolithic SBK	40%
Bronze Age	80%
Iron Age	70%
Greek	20%
Roman	60%

Combine classifications



Classify and combine



Each file has
37838 3D points ≈ 113.000 values
+ surface image

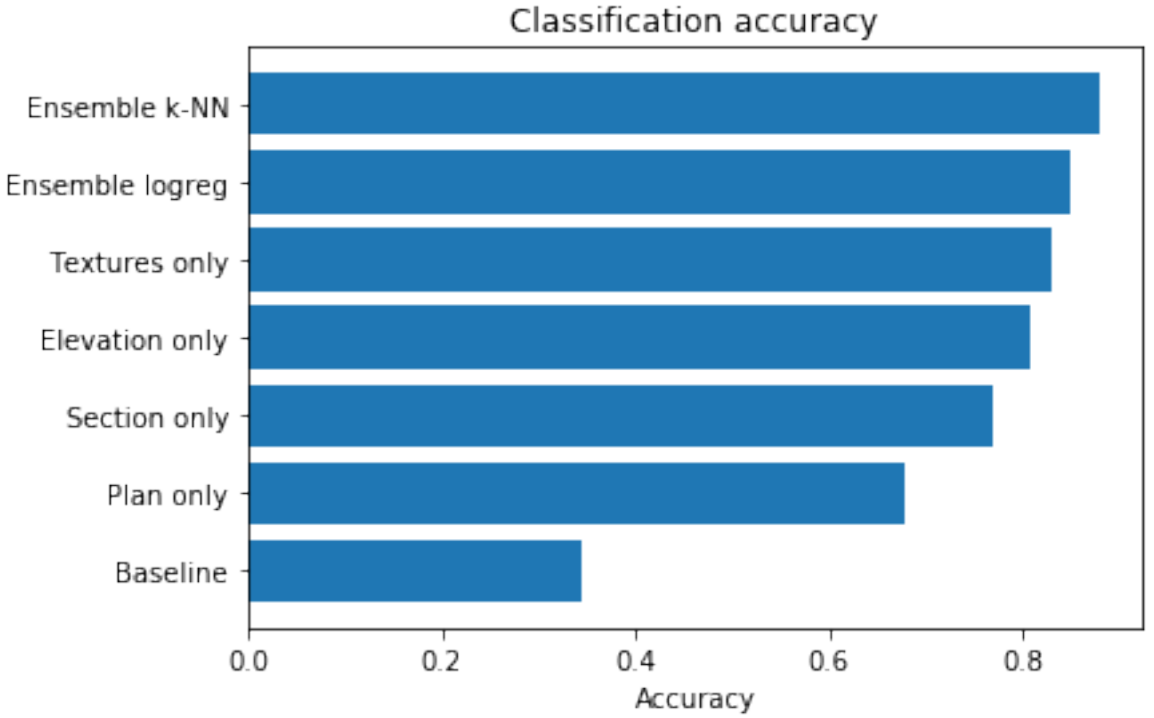
3 rendered views, 1 color
texture, 299x299 pixels
≈ 54.000 values

4 sets of 2048 features
≈ 8000 values

4 sets of 6 probabilities
= 24 values

Most fitting class: 1 value

Results: Accuracy



Results: confusion matrix

			Predictions					
			Bronze Age	Greek	Iron Age	Neolithic Linear Pottery Culture (LBK)	Neolithic Stroked Pottery culture (SBK)	Roman
True class		Bronze Age	97 %	0 %	0 %	3 %	0 %	0 %
		Greek	0 %	87 %	3 %	0 %	0 %	0 %
		Iron Age	7 %	0 %	83 %	0 %	0 %	0 %
		Neolithic Linear Pottery Culture (LBK)	0 %	0 %	0 %	100 %	0 %	0 %
		Neolithic Stroked Pottery culture (SBK)	10 %	0 %	0 %	40 %	50 %	0 %
		Roman	0 %	14 %	43 %	0 %	0 %	43 %

Evaluation

Practical results

- *Evaluation*
  Accuracy is not the only criterion

- *Transfer learning*
  Transfer learning does improve the results considerably, on small data sets.

- *Choice of model*
  Investigate several suitable alternatives, but stick with those you understand.

- *Apply suitable corrections to the results*
  Expect data to be incomplete, unbalanced or non-homogeneous, know how to handle these data.

What could be improved?

- *Choose a suitable scanning method for ML from the onset*
  While point clouds with tens of thousands points provide enough information,
  missing textures certainly are a problem. Therefore 3D-scanning methods that provide a texture should be preferred.

- The usual: Larger data set, faster hardware, more time...

https://github.com/aot29/pottery_classifier/