THE CYPRUS INSTITUTE
RESEARCH•TECHNOLOGY•INNOVATION

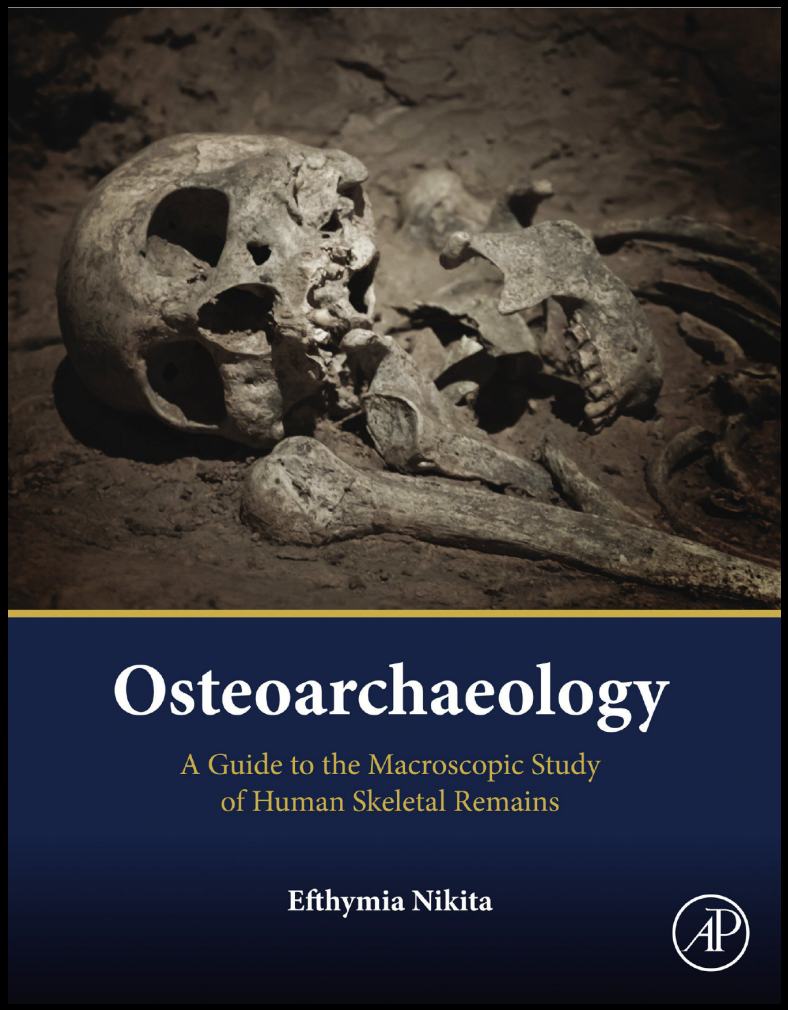
SexEst: An open access web application for metric skeletal sex estimation

Efthymia Nikita & Chrysovalantis Constantinou

Science and Technology in Archaeology and Culture Research Center, The Cyprus Institute

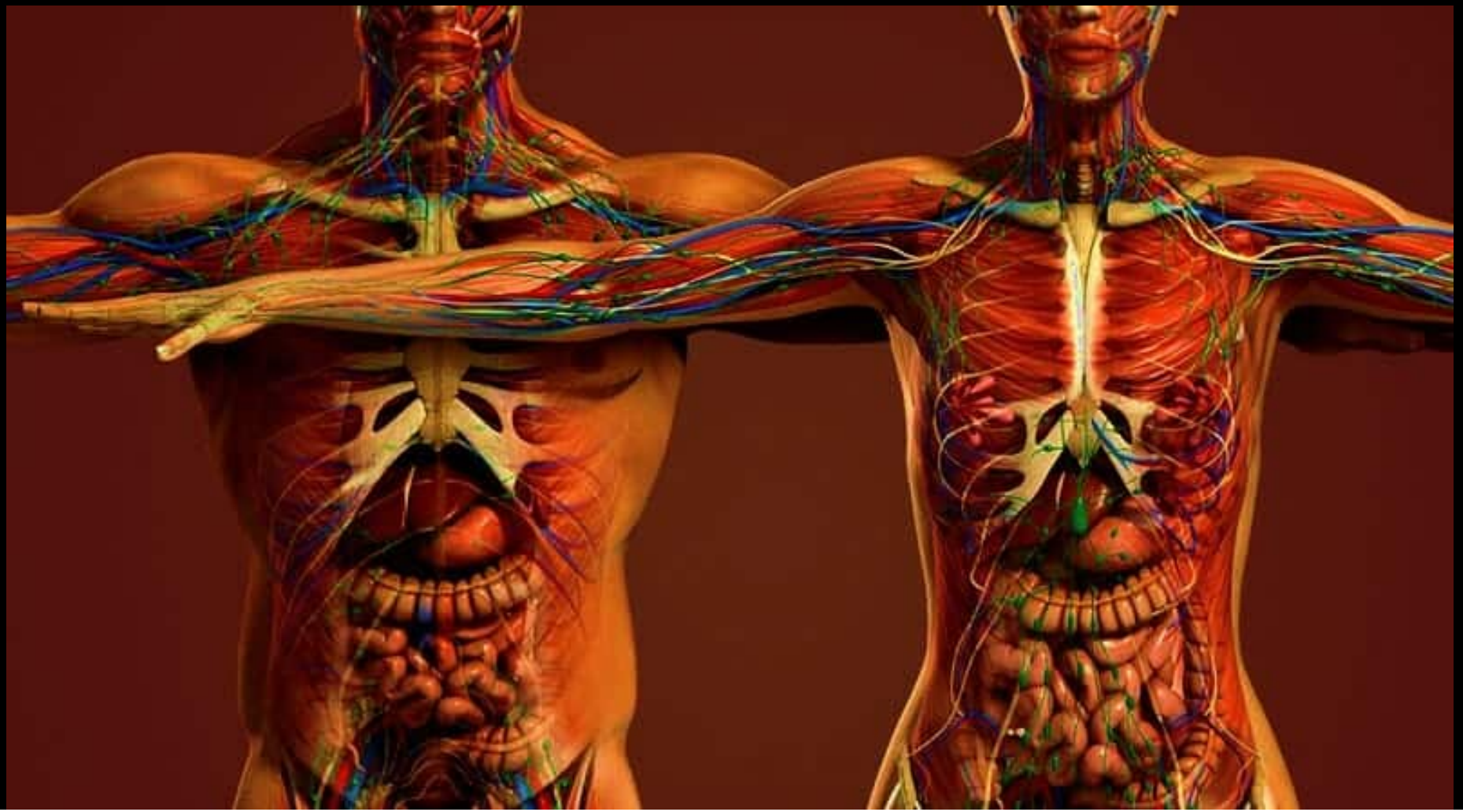
Computation-based Science and Technology Research Center, The Cyprus Institute

What is osteoarchaeology?

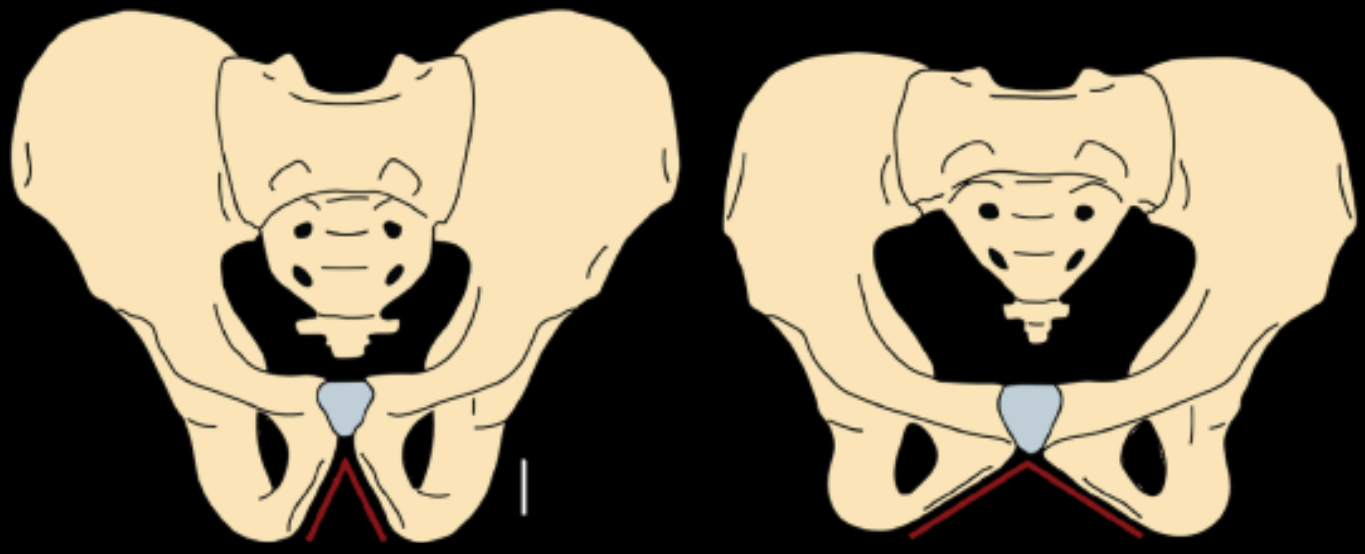


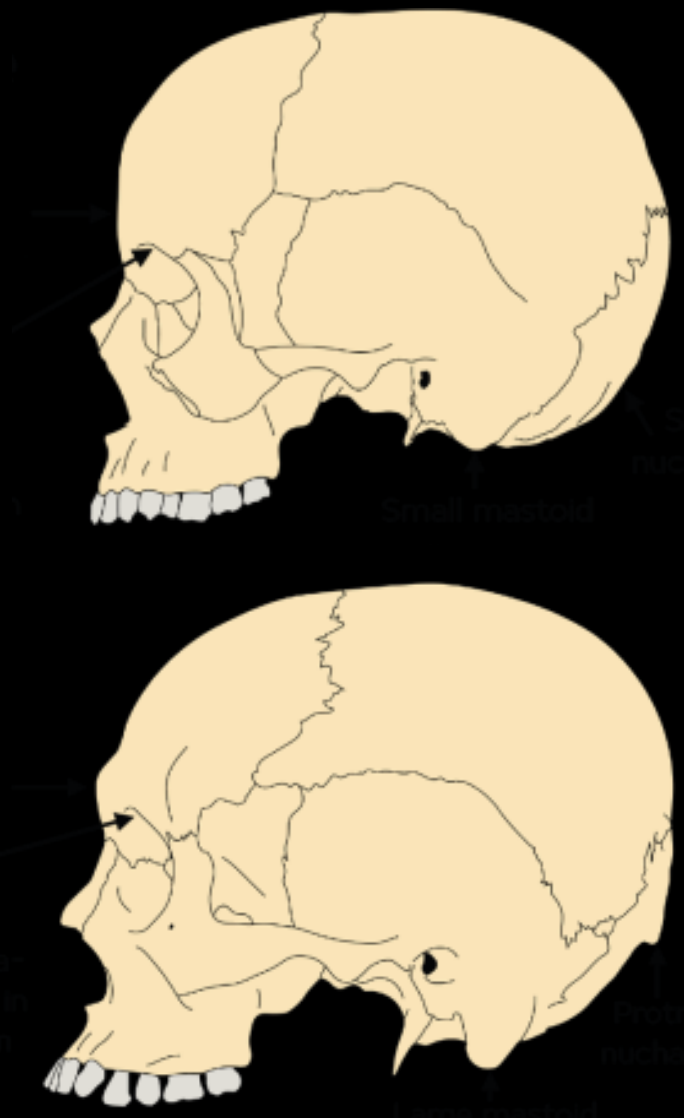
Study of human skeletal remains from archaeological contexts

Sex...even more important than you think!

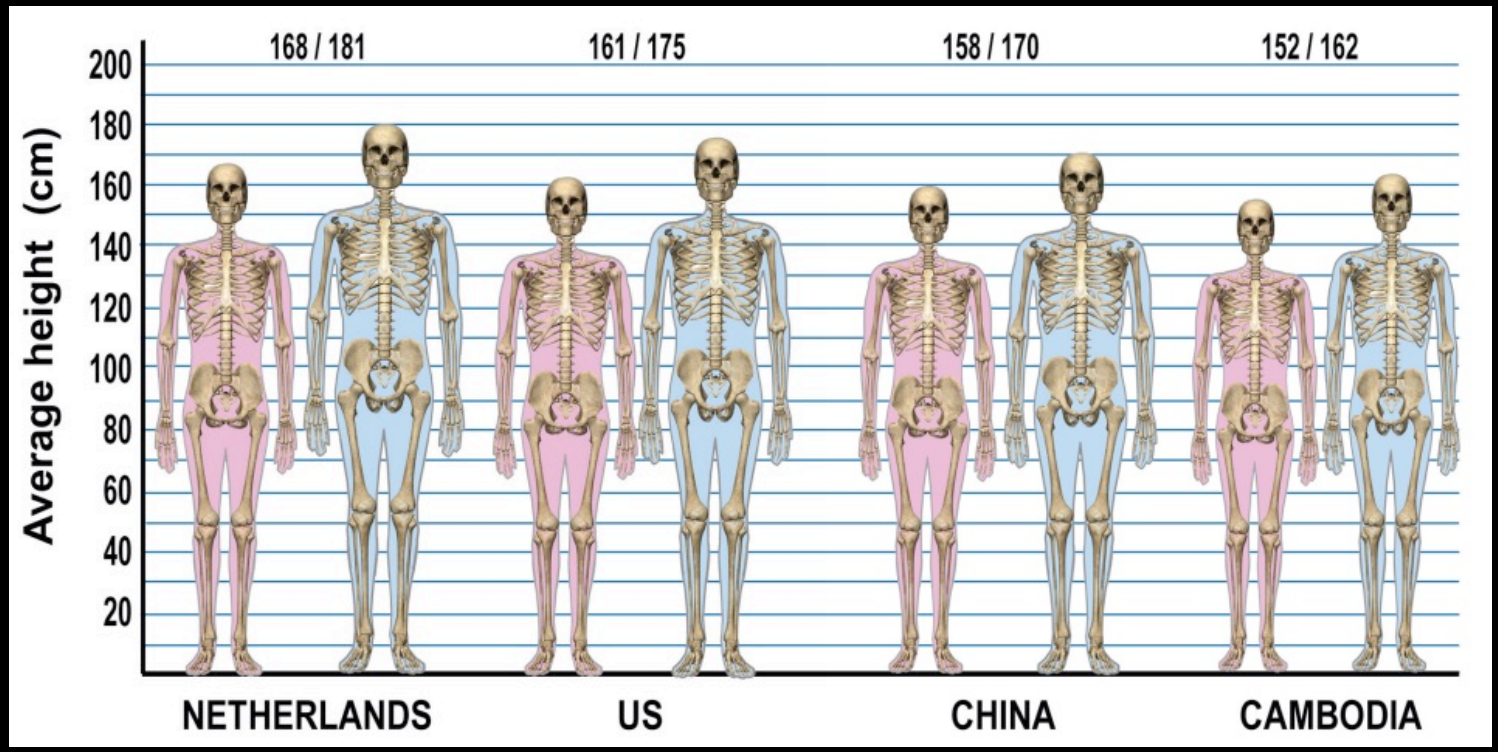


Morphological sex estimation





Metric sex estimation



Making life easier…

osteomics

https://osteomics.com/

Sex diagnosis

CalcTalus - Publication available

CADOES - Publication available

Ammer-Coelho - Publication available

SeuPF - Publication available

DSP (Depcreated, there is now a DSP2 by the official team)

Limitations of current models

- Based on modern assemblages (secular trends?)
- Population-specific methods

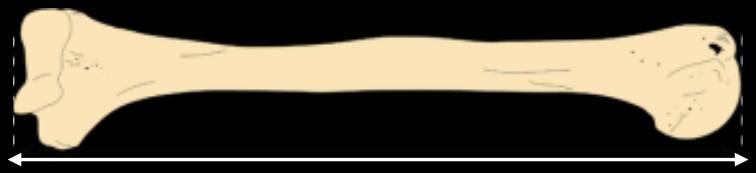
Aim

- Develop an open access and open source web application for multiregional sex prediction

Materials

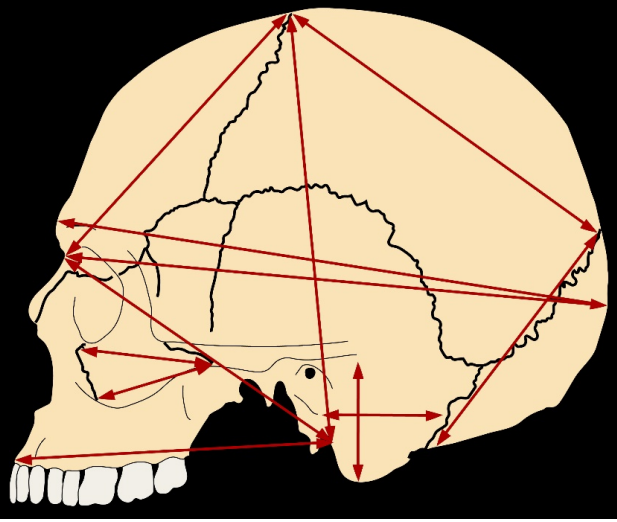
Goldman Dataset

- postcranial measurements
- 1538 skeletons
- from various geographic locations
- data collected by Dr. Benjamin Auerbach



William W. Howells Dataset

- cranial measurements
- 2524 skeletons
- from various geographic locations
- data collected by Dr. William W. Howells



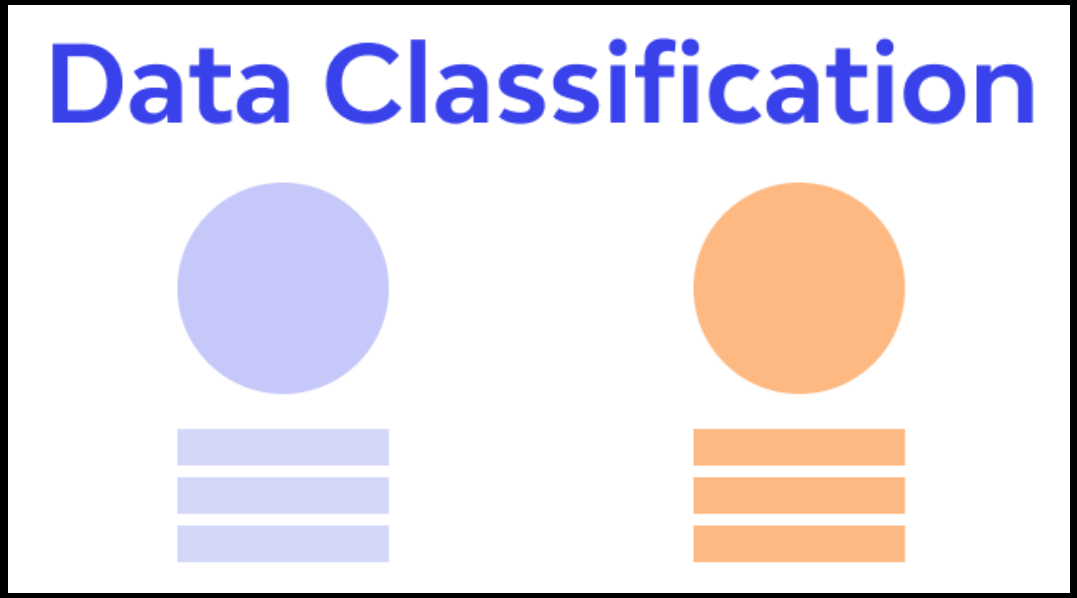
Different machine learning classifiers:

- Logistic Regression
- Decision Trees
- Support Vector Machine
- Gaussian Process
- Gradient Boosting
- Random Forest
- Ada Boosting
- Extra Trees
- Gaussian Naive Bayes
- k-Nearest Neighbors
- Linear Discriminant Analysis
- Quadratic Discriminant Analysis
- Extreme Gradient Boosting
- Light Gradient Boosting

Methods

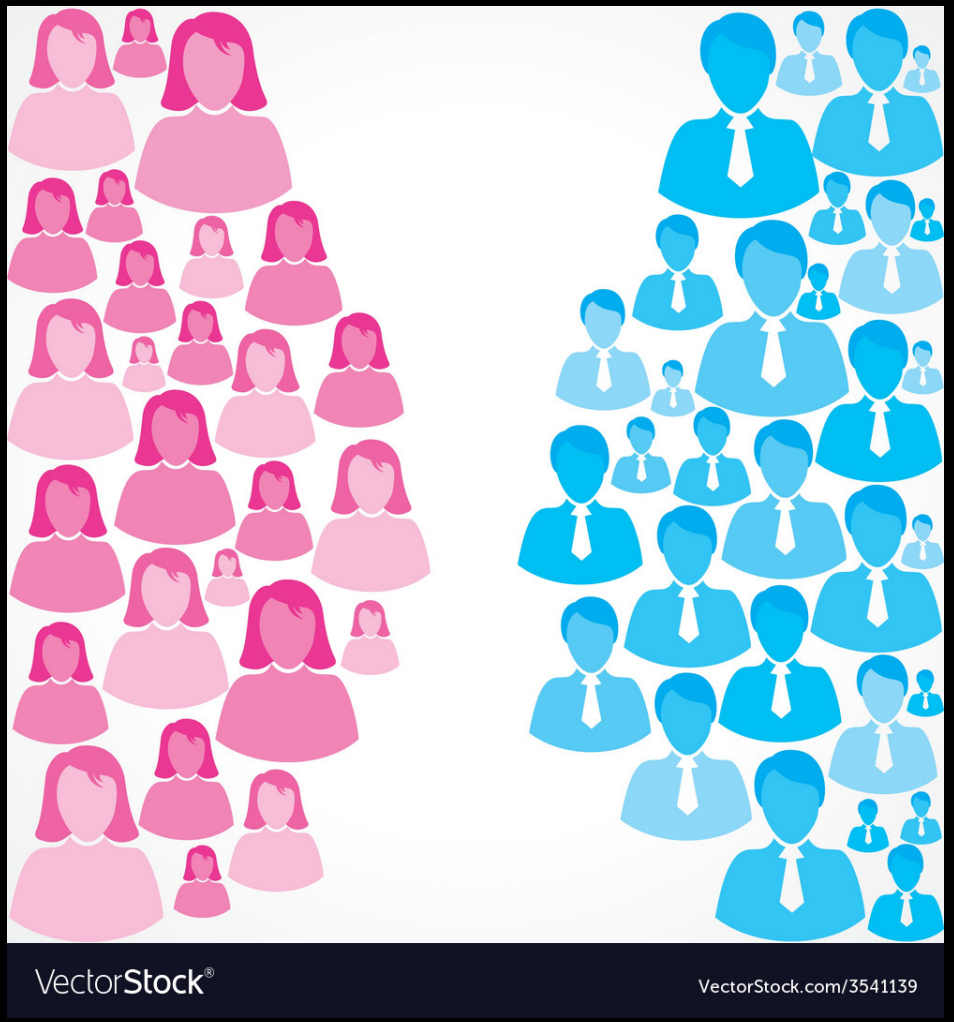
- Different combinations of variables
- Full datasets and missing values

- Python programming language



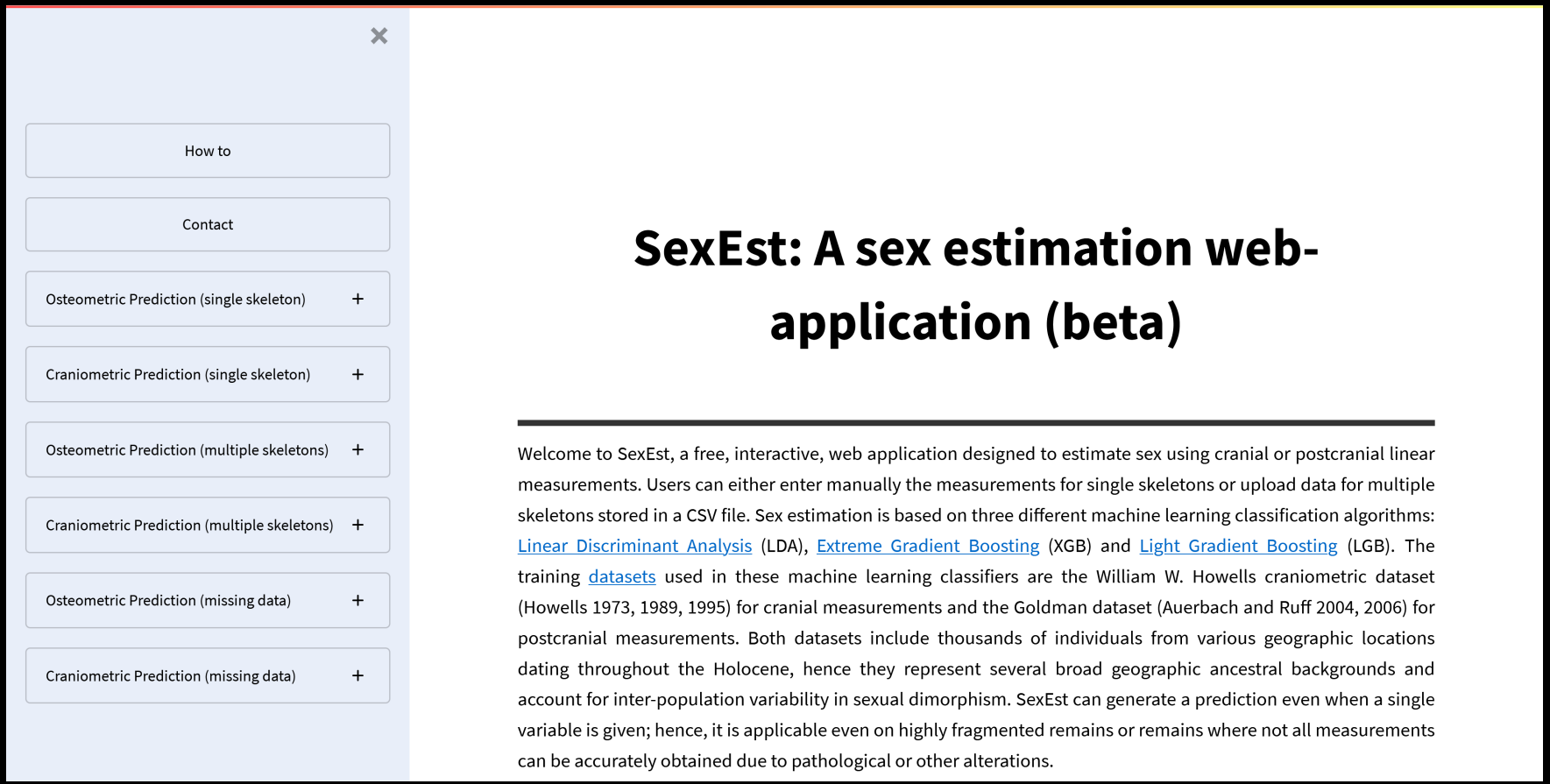
Results

- Highest accuracies
  - Extreme Gradient Boosting
  - Light Gradient Boosting
  - Linear Discriminant Analysis
  - no particular bias in classification
- Cranial measurements (>85%)
- Postcranial measurements (> 87%)



SexEst web application

http://sexest.cyi.ac.cy/



Osteometric prediction – single skeleton

SexEst: A sex estimation web-application (beta)

	BIB	HML	HHD	RML	FML	FBL	FHD	TML	FEB	TPB	HEB
0	266.0	299.0	42.0	255.0	423.0	413.0	46.8	341.0	75.0	69.0	61.2

The probability of the individual being male is 78.98% and the probability of being female is 21.02%

The model was trained using 1528 cases. The dataset was split into a training set and a test set with proportions 70% and 30%, respectively. The model was then trained using the training set and GridSearchCV, which optimized the model's hyperparameters and cross-validated it. The trained model was then tested using the test set, achieving an accuracy of 90.41%.

Osteometric prediction – multiple skeletons

Calculate

Craniometric Prediction (single skeleton) +

Osteometric Prediction (multiple skeletons) −

Choose a CSV file

Drag and drop file here

Limit 200MB per file • CSV

Browse files

sample_dataset_oste… ×

367.0B

Select a model

XGB

Calculate

Craniometric Prediction (multiple skeletons) +

	BIB	HML	HHD	RML	FML	FBL	FHD	TML	FEB	TPB	HEB
0	268.0	352.5	47.07	264.0	463.5	460.5	45.47	387.0	78.0	73.0	60.25
1	257.0	286.5	45.37	236.0	404.0	401.5	42.94	355.0	77.25	66.5	58.5
2	256.0	313.25	52.16	237.5	441.25	436.0	48.41	337.5	79.75	74.0	63.5
3	254.0	292.5	44.36	233.0	417.5	413.5	45.95	346.0	74.5	67.0	59.5
4	256.0	299.5	43.02	240.0	411.75	408.0	43.91	359.25	79.75	68.0	65.5

	Male	Female
0	99.77	0.23
1	98.2	1.8
2	99.68	0.32
3	98.07	1.93
4	99.92	0.08

Please note that any rows containing missing data will be dropped as these models have been optimized to work with all 11 variables submitted. You can try the **Osteometric Prediction (missing data)** mode for any cases/rows containing missing data.

The model was trained using 1528 cases. The dataset was split into a training set and a test set with proportions 70% and 30%, respectively. The model was then trained using the training set and GridSearchCV, which optimized the model's hyperparameters and cross-validated it. The trained model was then tested using the test set, achieving an accuracy of 90.41%.

Osteometric prediction – missing values

Next steps

Wiley Online Library

Search

International Journal of Osteoarchaeology

RESEARCH ARTICLE

SexEst: An open access web application for metric skeletal sex estimation

Chrysovalantis Constantinou, Efthymia Nikita

First published: 13 April 2022 | https://doi.org/10.1002/oa.3109

- Expansion of post-cranial variables
- Validation in different assemblages worldwide

Acknowledgements

Benjamin Auerback (open access datasets)

- NI4OS-Europe project (European Commission under Horizon 2020 European Research Infrastructures grant agreement no. 857645)
- People in Motion project (European Regional Development Fund and Republic of Cyprus through Research and Innovation Foundation; EXCELLENCE/1216/0023)
- Promised project (European Union's Horizon 2020 Research and Innovation Program under grant agreement no. 811068)

RESEARCH
& INNOVATION
FOUNDATION

Structural Funds
of the European Union in Cyprus