Collaborative Research in Machine Learning on the African Continent

- Gladys Kingori (Kenya)
- Musa A. M. Hussien (Sudan, South Africa)
- Damilare Babatunde (Nigeria)
- Georgies Alene Asres (Ethiopia)
- G. Solola (Nigeria)
- Paul K. Ngumbi (Kenya)
- Jean Baptiste Fankam Fankam (Cameroon, South Africa)
- Christian Aim´e Njeumen (Cameroon)
- Alhadji Malloum† (Cameroon)
- Ibrahim Isah (Nigeria)
- **Omololu Akin-Ojo (Nigeria; Previously: Rwanda) oakinojo@gmail.com**

UNIVERSITY OF IBADAN

East African Institute for Fundamental Research (EAIFR)

United Nations
Educational, Scientific and
Cultural Organization

ICTP - East African Institute
for Fundamental Research
under the auspices of UNESCO

The Abdus Salam
International Centre
for Theoretical Physics

Republic of Rwanda

Ministry of Education

UNIVERSITY of RWANDA

CMERIAUX@EAIFR.ORG || DIRECTOR@EAIFR.ORG

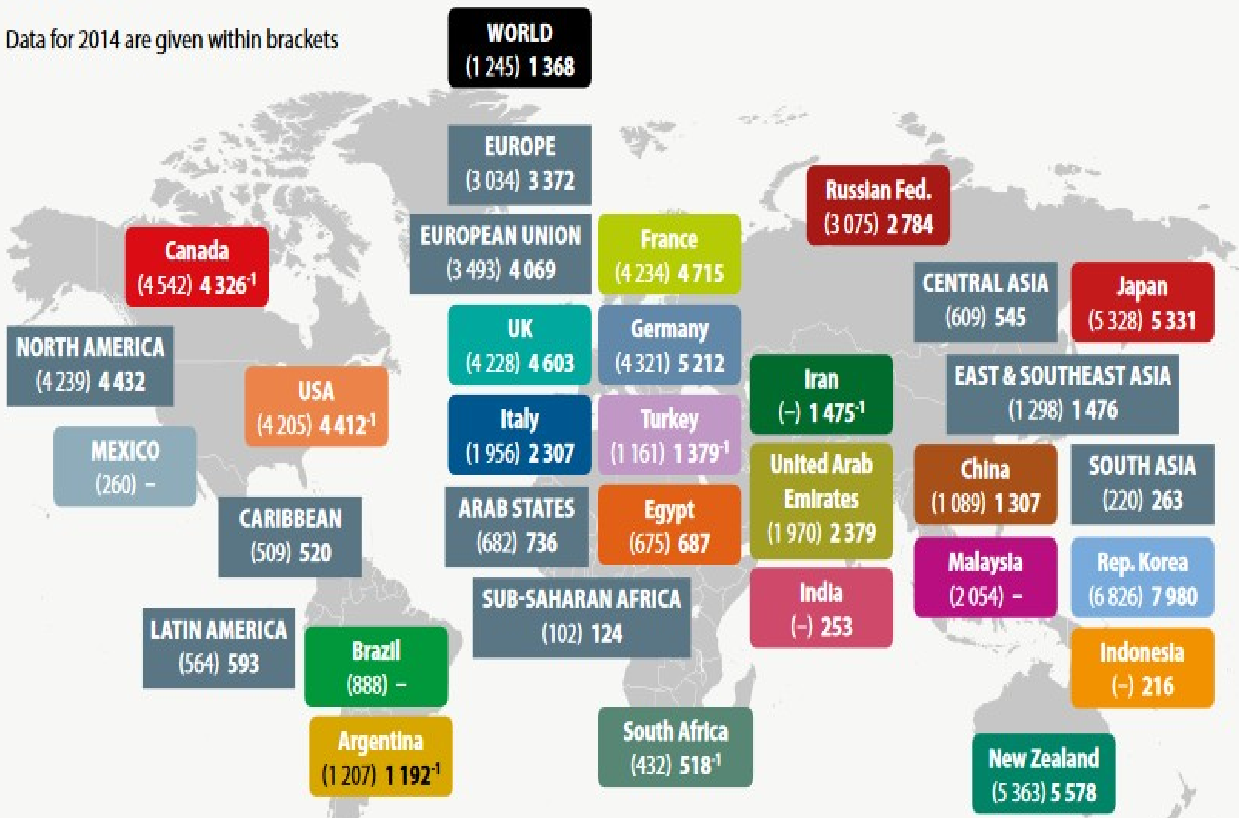
Africa = Poor?

Research can make Africa More Prosperous

Research and Science and Technology are paths to making Africa more prosperous. Unfortunately:

Figure 1.3: **Researchers (FTE) per million inhabitants, by region and selected country, 2014 and 2018**

Data for 2014 are given within brackets



13.7% Growth in the global number of researchers (FTE) between 2014 and 2018

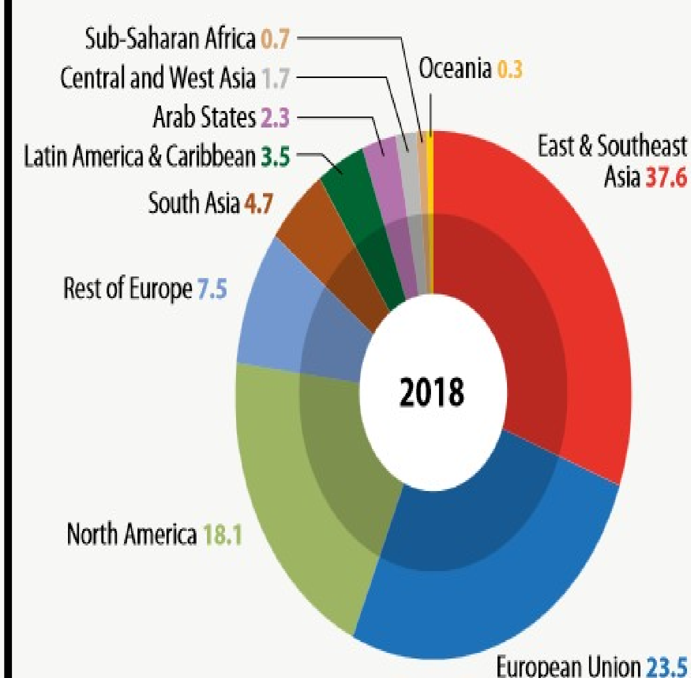
4.6% Growth in global population between 2014 and 2018

1.6% Growth in expenditure per researcher between 2014 and 2018

Source: global and regional estimates based on country-level data from the UNESCO Institute for Statistics, August 2020, without extrapolation

Figure 1.9: **Global trends in researchers (FTE)**

Global shares of researchers by region, 2015 and 2018 (%)



Increase Intellectual Capacity in Africa

1. Increase number of researchers in Africa
2. Increase level/quality of research
3. Increase level/quality of teaching/learning

Three Hurdles of Researchers in Africa

(i) Excessive Teaching and Administrative loads

(ii) Isolation: Lack of Proper Mentoring in research
Limited access to Journal Papers

(iii) Limited Resources: * Computational
* Experimental
* Journal Papers

Possible Solutions

(i) Sharing Teaching Resources Online & engaging volunteer visiting lecturers

(ii) Mentoring: Pairing African Researchers with Excellent researchers in the developed world

(iii) Sharing available resources within Africa and from outside Africa e.g., Joint Computational Facilities

CRAMM/MLESMD

Collaborative Research in Atomistic and Molecular Modeling (CRAMM)

1. Joint weekly meetings

(a) Forces participants to make/take time for research

(b) Reduces isolation

2. Sharing Resources

(a) Computational Resources

(b) Human Resources

(c) Ideas

(d) Joint training

MLESMD

Machine Learning in Electronic Structure and Molecular Dynamics

ICTP

16 - 20 May 2022
Kigali - Rwanda

Further information:
http://indico.ictp.it/event/9787/
smr3708@ictp.it

This is the culminating workshop of the Machine Learning (ML) in Electronic Structure (ES) and Molecular Dynamics (MD) training workshop that began in October 2021. The focus is to train African researchers in ML and its application in ES and MD.

Description:

A recap of the Python language, ML and Materials Databases will be made. This will be quickly followed by joint projects in materials discovery. Each day will feature hands on training sessions as well as two to three talks from experts in the field presenting their research.

Topics:

- Introduction to the Python language
- Introduction to Machine Learning
- Databases for Materials Discovery
- Machine Learning for force fields development for
- Molecular Dynamics Simulations
- Machine Learning in Electronic Structure

Directors:

O. Akin-Ojo (ICTP-EAIFR & Univ. Rwanda)
A. Hassanali (ICTP, Trieste, Italy)
A. Laio (SISSA, Trieste, Italy)
S. Ndengue (ICTP-EAIFR & Univ. Rwanda)
G-M. Rignanese (UCL, Belgium)
M. Scheffler (Fritz-Haber MPG, Germany)

Local Organisers:

O. Akin-Ojo (ICTP-EAIFR & Univ. Rwanda)
S. Ndengue (ICTP-EAIFR & Univ. Rwanda)

ICTP Scientific Contact:

A. Hassanali (ICTP, Trieste, Italy)

Speakers:

O. Akin-Ojo (ICTP-EAIFR & Univ. Rwanda)
J. Behler (Univ. Gottingen, Germany)
M. Ceriotti (EPFL, Switzerland)
B. Cheng (IST, Austria)
G. Csanyi (Univ. Cambridge, U.K.)
S. Curtarolo (Duke Univ. U.S.A.)
S. de Gironcoli (SISSA, Trieste, Italy)
A. Jain (LBNL, USA)
E. Kucukbenli (Boston Univ., USA)
A. Laio (SISSA, Italy)
M. Parrinello (ETH, Zurich)
G-M. Rignanese (UCL, Belgium)
M. Scheffler (Fritz-Haber MPG, Germany
T. Smidt (UC Berkely, U.S.A.)

How to apply:

Online application:
http://indico.ictp.it/event/9787/
Female scientists are encouraged to apply.
Priority will be given to those who actively participated in the ML training from October 2021.

Grants:

There is no registration fee.

Deadline:

15 April 2022

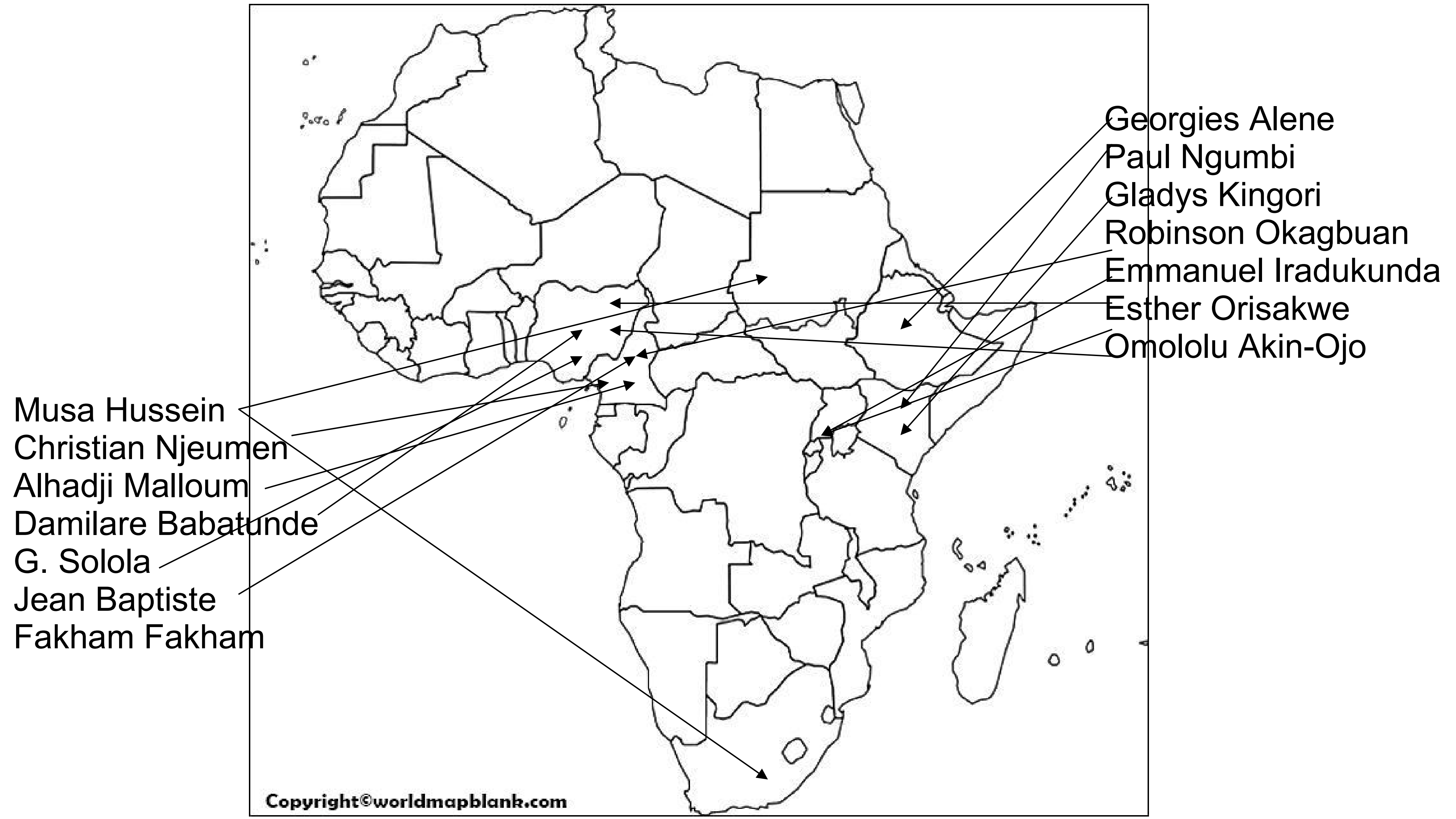
EAIFR

MOHAMMED VI UNIVERSITY

EPFL

ICTP International Centre for Theoretical Physics
www.ictp.it

Our TEAM


Musa Hussein
Christian Njeumen
Alhadji Malloum
Damilare Babatunde
G. Solola
Jean Baptiste
Fakham Fakham
Georgies Alene
Paul Ngumbi
Gladys Kingori
Robinson Okagbuan
Emmanuel Iradukunda
Esther Orisakwe
Omololu Akin-Ojo
Copyright©worldmapblank.com

OVERVIEW

- ***Material: Insulator or Conductor?***
- ***Bandgap: Direct or Indirect?***
- ***Search for High Thermal Conductivity materials***
- ***Search for Super-Hard materials***

Insulator or Conductor?

- **PbS: Metal or Insulator?**

  - Electrical Property from *Chemical Formula*

- **DATA:**

  - SNUMAT (Hybrid DFT Bandgaps. 18K Compounds)

- **Features/Descriptors**

  - Electronegativity, Mendeleev Number, Atomic weight, # of s, p, d, f valence, etc

- **Models**
  - ***Random Forest (RF) classifier***
  - ***LGBM classifier***
  - ***XGB classifier***

Insulator or Conductor?

PERFORMANCE

Figure 1: Models Performance

Models	RF	LGBM	XGB
Accuracy	0.921	0.921	0.923

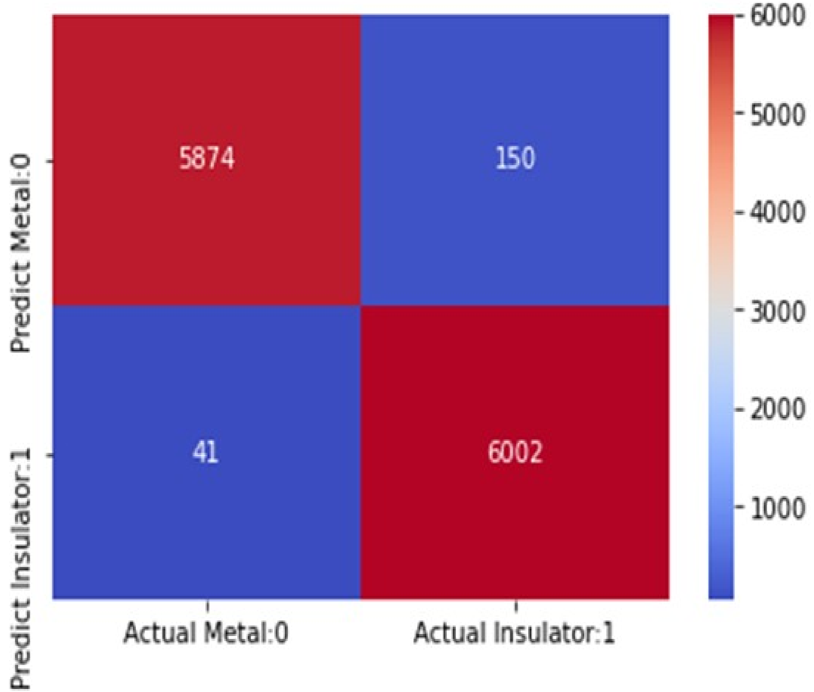
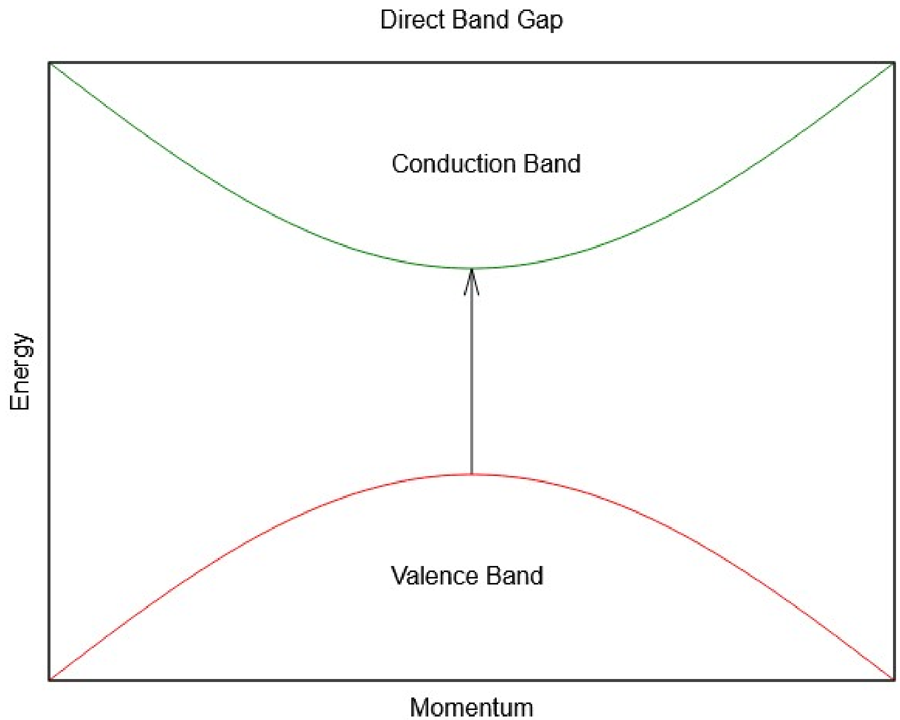
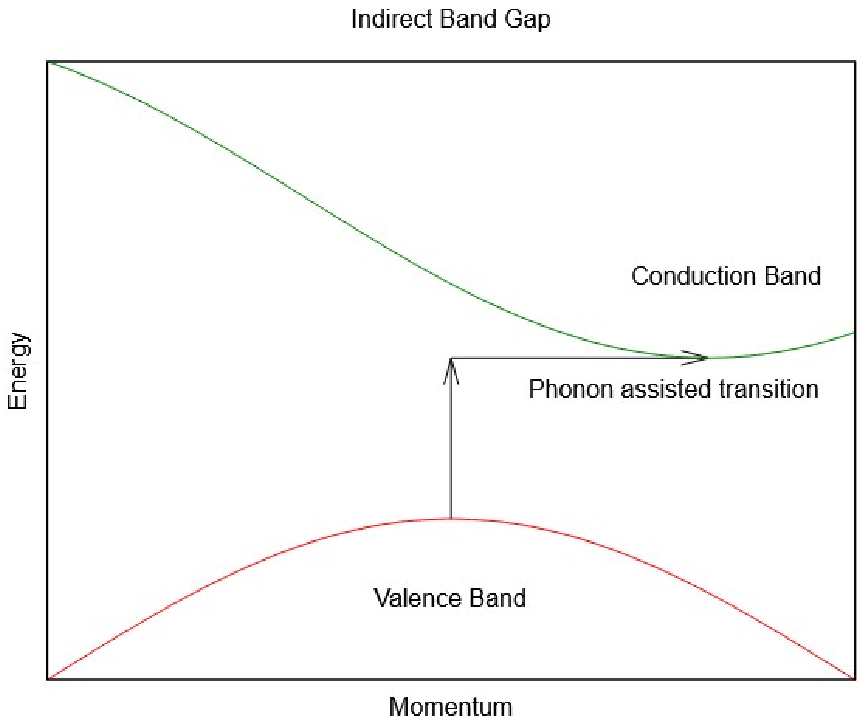


Figure 1: XGBClassifier Confusion matrix

BandGap: Direct vs. Indirect





BandGap: Direct vs. Indirect

- **Silicon: Direct or Indirect?**

  - Electrical Property from *Chemical Formula*

  - Electrical Property from Chemical Formula & Structure

- **DATA:**

  - SNUMAT (Hybrid DFT Bandgaps. 11K Compounds)

- **Features/Descriptors**

  - Electronegativity, Mendeleev Number, Atomic weight, # of s, p, d, f valence, (average) volume per atom, packing efficiency, density, etc

- **Models**
  - ***Random Forest (RF) classifier***
  - ***Extra Tree classifier***

BandGap: Direct vs. Indirect

PERFORMANCE

Table 2: Models Performance

Models	**RF**	**Extra Tree**
Accuracy	0.81	0.80

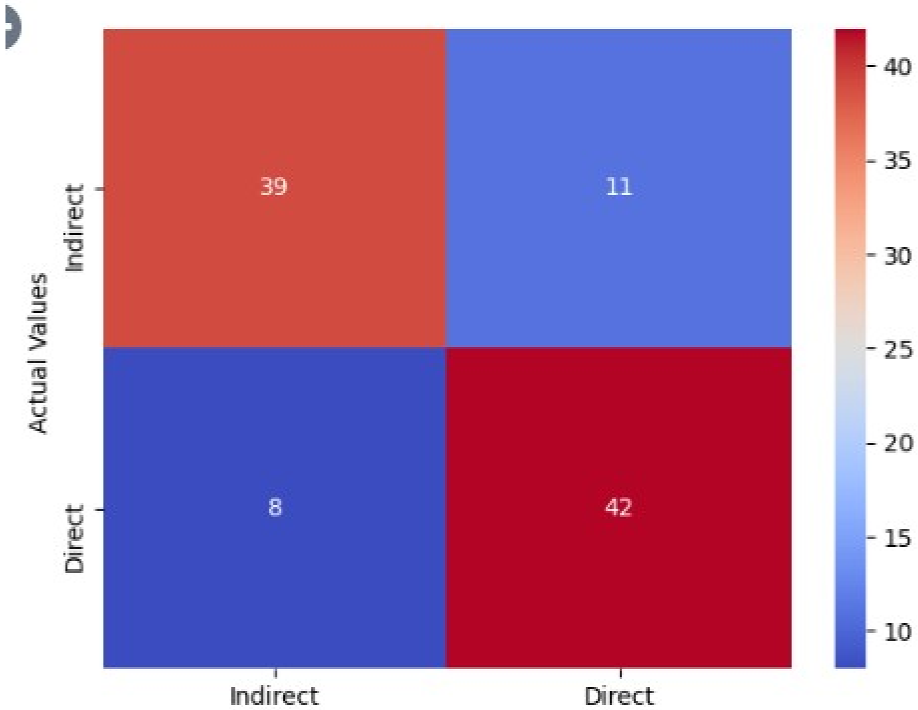


Figure 3: Confusion matrix for RF

Search: High K Materials

$$\vec{J}^Q = -\kappa \nabla T$$

- Thermal Property from Chemical Formula & Structure

- **DATA:**

  - AFLOW Database (5.5K Compounds)

- **Features/Descriptors**

  - Electronegativity, Mendeleev Number, Atomic weight, # of s, p, d, f valence, (average) volume per atom, packing efficiency, density, etc

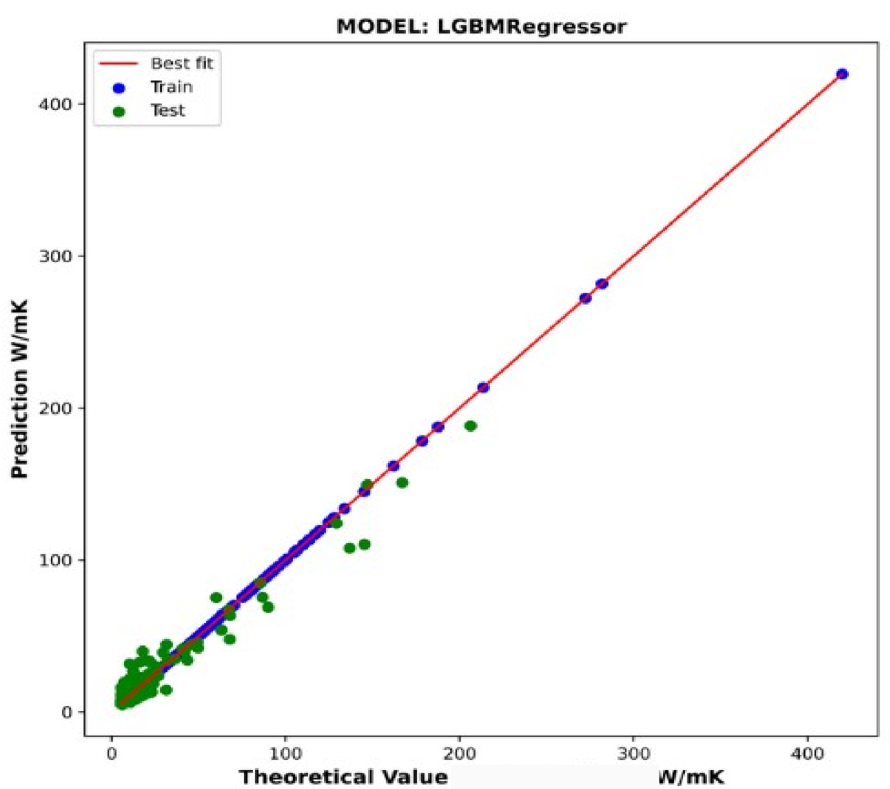
- **Models**

  - *HistGradient Boost Regression*
  - *Extra Tree Regression*

Search: High K Materials

PERFORMANCE

Models	R^2	MAE (W/mK)	RMSE
HistGradientBoosting	0.966	3.149	5.086
ExtraTree	0.947	3.727	6.326



HistGradientBoosting Regressor Model

Search: Hard Materials

- Bulk (B), Shear (G), Young (E) moduli, Possion ratio (ν)

  - Mechanical Properties from *Chemical Formula & Structure*

- DATA:

  - AFLOW Database (DFT 5.5K Compounds)

- Features/Descriptors

  - Electronegativity, Mendeleev Number, Atomic weight, # of s, p, d, f valence, (average) volume per atom, packing efficiency, density, etc

- Model
  - *Histogram Gradient Boosting Regressor*

Search: Hard Materials

PERFORMANCE

Table 4: Model Performance (HGBR)

Target	R^2	MEA	RMSE
B	0.980	6.11	9.664
G	0.931	8.00	12.965
v	0.955	0.0443	0.0365

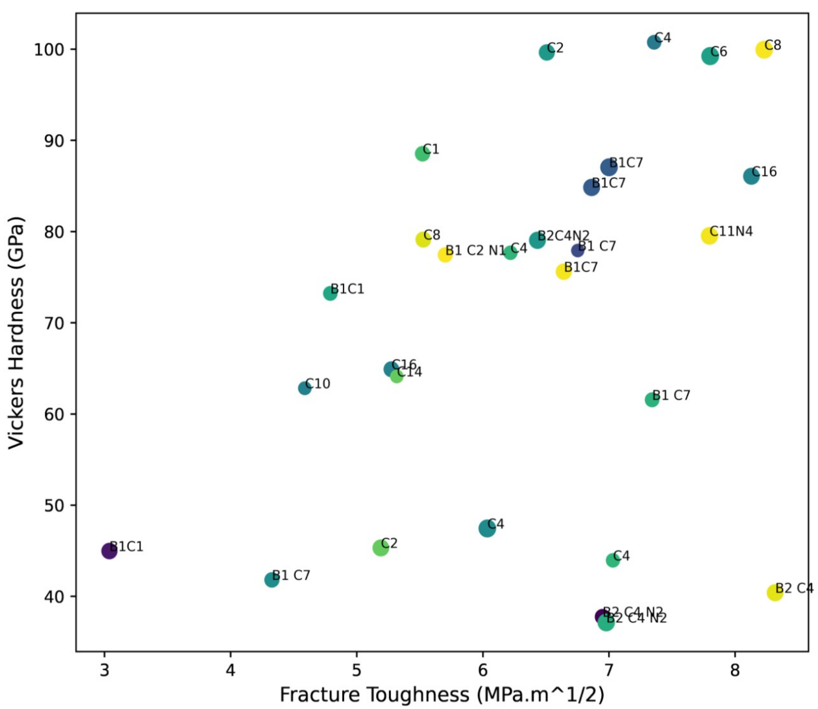
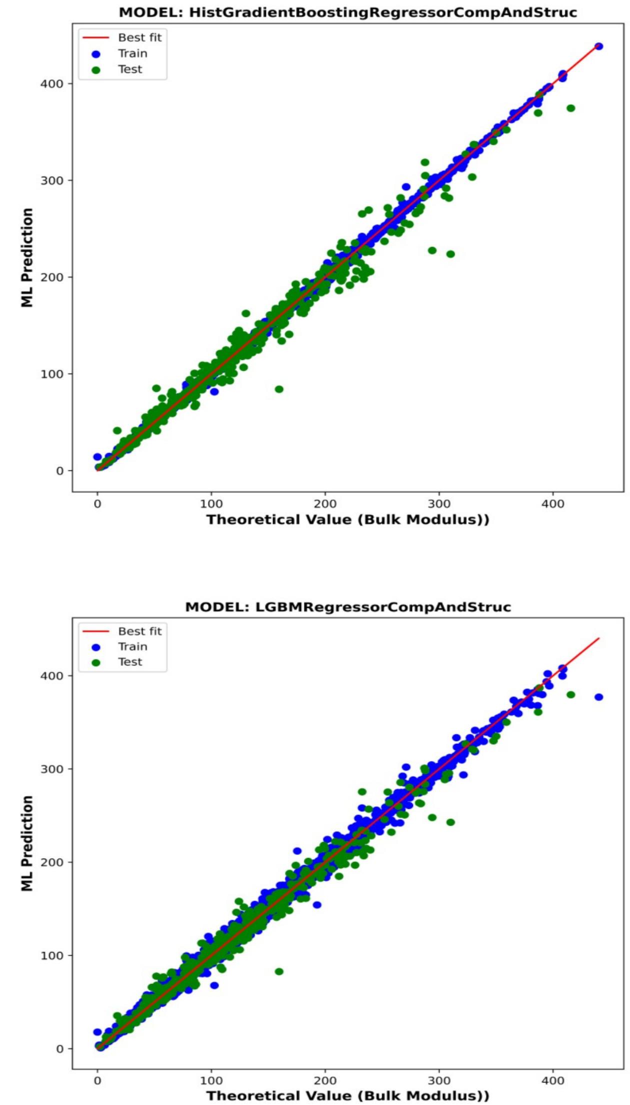


Figure: Hardness of Some materials



Bulk and Shear Modulus (HGBR)

Conclusion

- Work TOGETHER
- Research TOGETHER
- ML TOGETHER
- Science:
  - **Metal/Insulator**
  - **Direct/Indirect**
  - **High κ materials**
  - **Hard(er) Materials**

Acknowledgement

- Gladys Kingori (Kenya)
- Musa A. M. Hussien (Sudan, South Africa)
- Damilare Babatunde (Nigeria)
- Georgies Alene Asres (Ethiopia)
- Paul K. Ngumbi (Kenya)
- G. Solola (Nigeria)
- Jean Baptiste Fankam Fankam (Cameroon, South Africa)
- Christian Aim´e Njeumen (Cameroon)
- Alhadji Malloum† (Cameroon)
- Ibrahim Isah (Nigeria)
- Gian-Marco Rignanese (UCL Belgium)
- Govt. Of Rwanda (and UR)
- ICTP
- EAIFR

THANK YOU