

# MACHINE LEARNING COMPETITIONS: The Outlook from Africa

Dina Machuve, PhD  
CiML 2019  
13<sup>th</sup> December, 2019

# Dina Machuve

- Lecturer at Nelson Mandela African Institution of Science and Technology in Tanzania
- Co-organize Data Science Africa (DSA)
- PhD in Information and Communication Science and Engineering



# African Situation

- A huge focus in currently on capacity building



DEEP LEARNING  
INDABA

---

ZIND!



# Zindi Africa



- A social enterprise building the data science ecosystem in Africa.
- Data Science competition platform in Africa.
- Uses African datasets to solve African problems.
- Committed to solve Africa's toughest challenges

# Zindi Africa



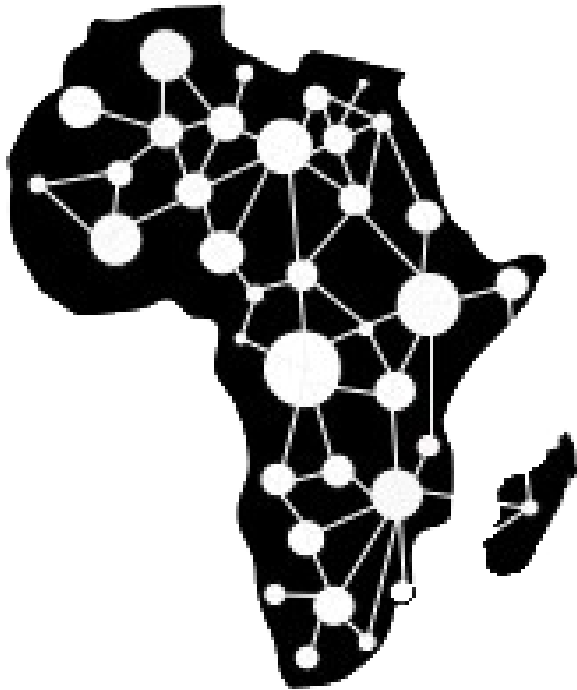
- 7000+ data scientists registered on Zindi
- 15% of Zindi users are non-African
- Challenge themes:
  - financial inclusion
  - transportation and mobility,
  - conservation,
  - agriculture,
  - fraud detection,
  - computer vision,
  - natural language processing

# Zindi Africa



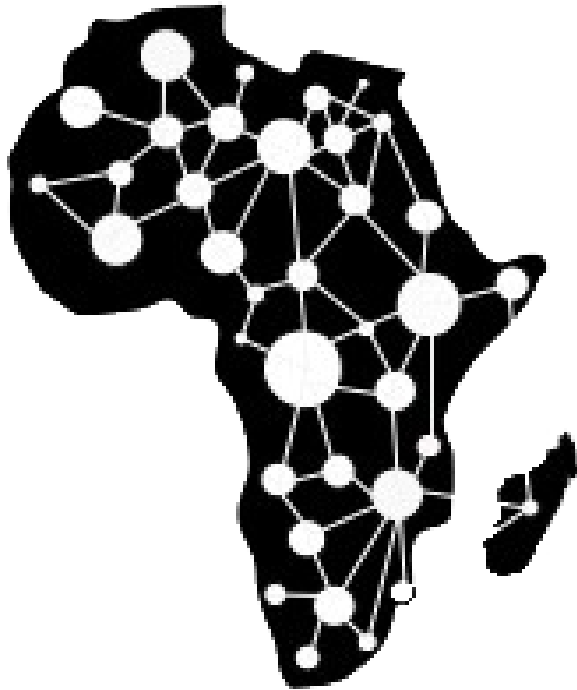
- Zindi outreach to users in Africa:
  - data science, AI, and ML meet-up groups
  - university clubs
  - Zindi Ambassadors
- Nigeria, South Africa, Kenya and Tunisia lead in challenges.

# Data Science Africa (DSA)



- A bottom-up capacity building initiative of African experts
  - Technical skills
  - Organizational skills
- Supplemented by international support
- Annual summer school and workshop since 2015
- Next DSA in Kampala, Uganda  
22<sup>nd</sup> -26<sup>th</sup> June, 2020

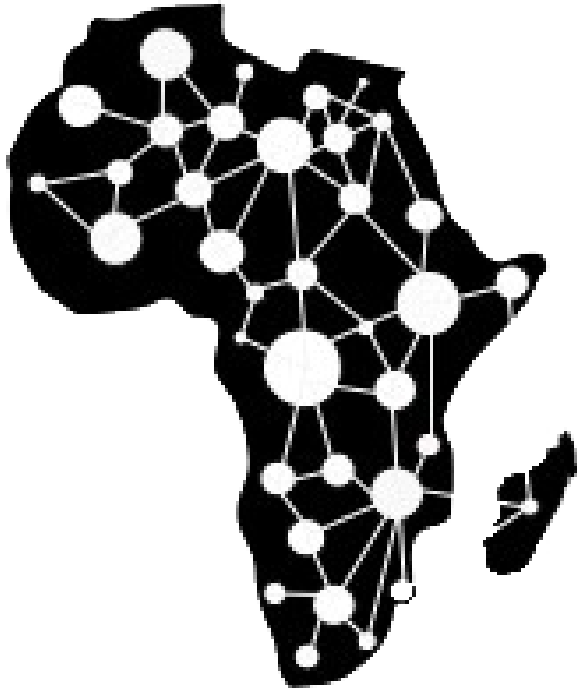
# Data Science Africa



- DSA Meetings
  - 3 days data science school
  - 2 days workshop
- Thematic areas:
  - Agriculture
  - Transport & infrastructure
  - Population health
  - Wildlife

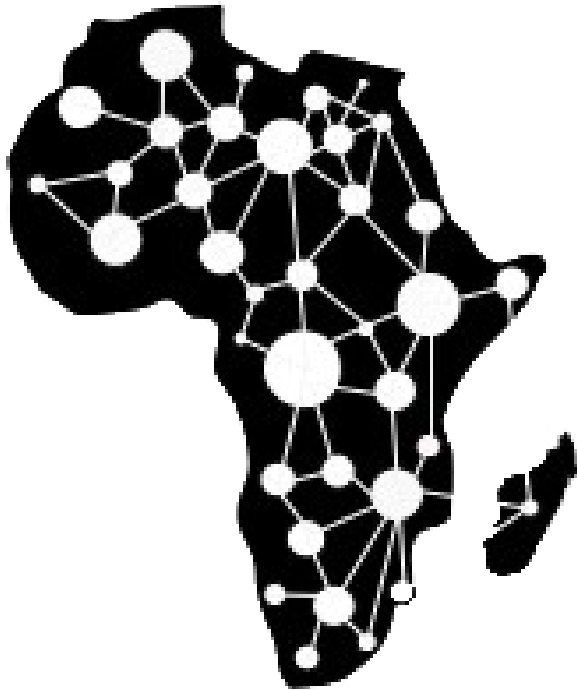


# Competitions at DSA



- Introduced in 2018
  - Entry to summer school
- In 2019 on CODALAB
  - Dataset from Uganda
  - Low participation before DSA
  - After a tutorial at DSA, participation improved

# End-to-end data science



- End-to-end data science
  - robust and sustainable local data science communities
  - focus on challenges that arise in the African context
  - deliver systems with actionable insights by end users

# End-to-end data science applications



(a) Healthy

(b) CBB



(c) CGM

(d) CMD

(e) CBSD

- Cassava Disease Classification
  - Challenge on Kaggle April 25<sup>th</sup> to June 1<sup>st</sup> , 2019
  - A real life dataset on a problem affecting food security.

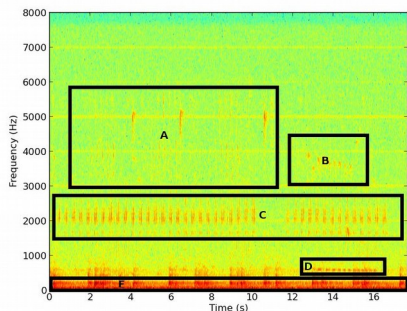
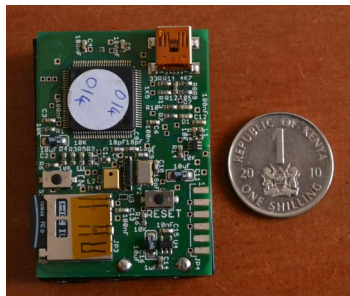
# End-to-end data science applications



Source: [CNN Health](#)

- Automated diseases diagnostics in Africa
  - Microscopy-Based Point of Care Diagnostics in Uganda (Quinn et al., 2016)
  - Diagnosis of malaria in thick blood smears
  - Tuberculosis in sputum samples
  - Intestinal parasite eggs in stool samples
  - Method used: Deep Convolution Neural Networks
  - Deployment: Smart-phone microscope adapter

# End-to-end data science applications



Data Paper

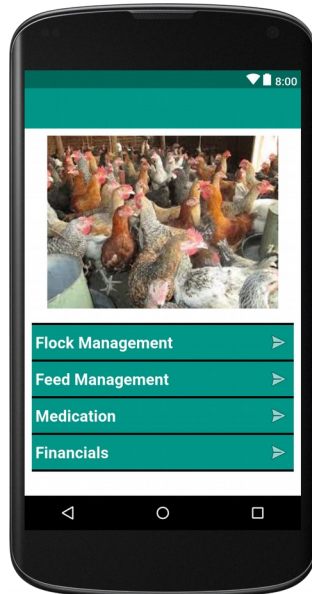
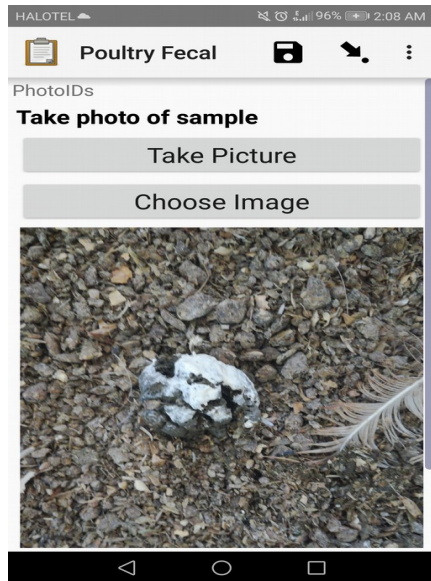
Biodiversity Data Journal 4: e9906  
<https://doi.org/10.3897/BDJ.4.e9906> (05 Oct 2016)

## A Bioacoustic Record of a Conservancy in the Mount Kenya Ecosystem

▼ Ciira wa Maina, David Muchiri, Peter Njoroge

- The Kenya Bioacoustics Project
  - Acoustic monitoring of ecosystems
  - Conservation efforts with limited budget
  - Focus on bird species which serve as **indicator species**
  - Inspired by MLSP 2013 Bird Classification Challenge
  - Acoustic data – 2700 recordings, 300 annotated, 36 bird species identified

# End-to-end data science applications



- Automated **Poultry** diseases diagnostics in Africa
  - early detection of Coccidiosis, Salmonella and Newcastle poultry diseases from fecal samples
  - enabled by AI
  - will save time for testing the diseases to about few hours
  - conventional methods of producing culture take over a week for diagnosis

# Deep Learning Indaba



**DEEP LEARNING  
INDABA**

---

- Strengthen African Machine Learning and AI
  - Build communities
  - Create leadership
  - Recognize excellence and innovation across the continent
- Annual gathering of the African AI community, since 2017

# Deep Learning Indaba



**DEEP LEARNING  
INDABA**

---

- Hackathon Track at DLI 2019
  - Mainly tutorial work
  - Engaged participation thereafter

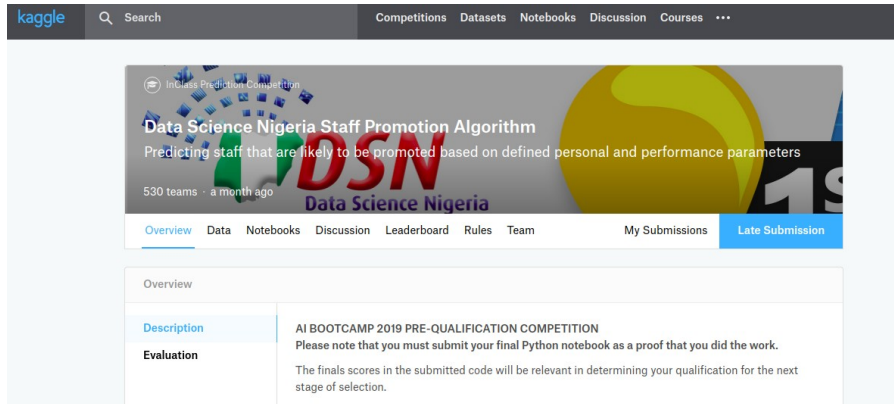


# Data Science Nigeria



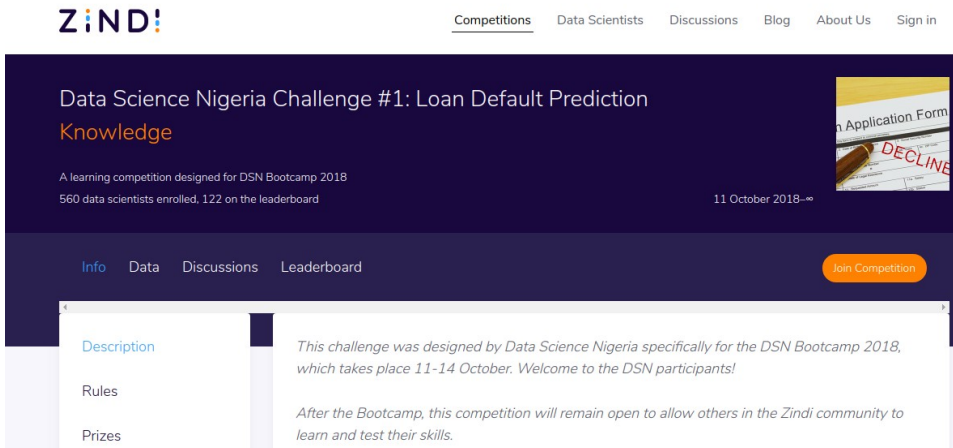
- Building 1 million African AI talents in 10 years
- Not-for-profit organization on local capacity building in AI since 2017

# Data Science Nigeria



The screenshot shows the Kaggle interface for a competition titled "Data Science Nigeria Staff Promotion Algorithm". The main heading is "Data Science Nigeria Staff Promotion Algorithm" with a subtitle "Predicting staff that are likely to be promoted based on defined personal and performance parameters". It indicates "530 teams · a month ago". The navigation bar includes "Overview", "Data", "Notebooks", "Discussion", "Leaderboard", "Rules", "Team", "My Submissions", and "Late Submission". The "Description" section states: "AI BOOTCAMP 2019 PRE-QUALIFICATION COMPETITION. Please note that you must submit your final Python notebook as a proof that you did the work." The "Evaluation" section states: "The finals scores in the submitted code will be relevant in determining your qualification for the next stage of selection."

- Kaggle & Zindi Competitions
  - drives capacity building through competitive engagement
  - Dataset focused on addressing local business or social issues
  - Entry criteria to ML bootcamp and mentoring



The screenshot shows the Zindi interface for a competition titled "Data Science Nigeria Challenge #1: Loan Default Prediction". The main heading is "Data Science Nigeria Challenge #1: Loan Default Prediction" with a subtitle "Knowledge". It indicates "A learning competition designed for DSN Bootcamp 2018" and "560 data scientists enrolled, 122 on the leaderboard". The date is "11 October 2018". The navigation bar includes "Info", "Data", "Discussions", "Leaderboard", and a "Join Competition" button. The "Description" section states: "This challenge was designed by Data Science Nigeria specifically for the DSN Bootcamp 2018, which takes place 11-14 October. Welcome to the DSN participants!" The "Rules" section states: "After the Bootcamp, this competition will remain open to allow others in the Zindi community to learn and test their skills."

# Nairobi Chapter WiMLDS



- A community for women (and men) with interests or working in data science since 2017
  - 2,778 followers
  - Learning from world-class AI researchers, engineers and practitioners
  - Strengthening data science skills
  - Community champions

# Nairobi Chapter WiMLDS



- Competitions are regular on meetup sessions
  - Local challenges on Github
  - Use past challenges on Kaggle and Zindi on several topics

# WeRobotics



- Not-for-profit organization focusing on
  - Sustainable development
  - Humanitarian aid
  - Public health
  - Environmental issues

# WeRobotics



- Local presence through Flying Labs
- Two challenges:
  - Open AI Challenge  
Tanzania, August 2018
  - The Unusual Solvers  
Competition: \$200K + in prizes in June, 2019

# WeRobotics



- Outcomes from Open AI Tanzania Challenge:
  - Total of 153 individuals signed up
  - 44 were from a country in Africa
  - Total of two winners, one from Austria and the other from US
  - Providing free AWS credits as part of the challenge would have increased participation of African AI developers

# African Master's in Machine Intelligence



- First cohort in 2018
- Applications for AMMI
  - 2018 a total of 1000
  - 2019 a total of 7053
  - Enrollment 35 per cohort



# Concluding remarks

- There is high potential of datasets for Africa
- The datasets are focused on real world problems
- Call for collaboration:
  - to develop more challenges in African context using the datasets
  - curating the datasets to fit on competition platforms
  - contribute to build capacity for African AI developers to participate in more challenges

“The most urgent task facing Africa today, and for decades ahead, is creating opportunities for this next generation.” - Barack Obama

**Thank You**