Policy Brief

EVALUATING GENDER BIAS IN AI APPLICATIONS USING HOUSEHOLD SURVEY DATA

Seth Goodman, Katherine Nolan, Rachel Sayers, Ariel BenYishay, Jacob Hall, Mavis Zupork Dome & Edem Selormey

In Brief

Over the past year, Aiddata and CDD-Ghana have been working on a gender-related artificial intelligence (AI) project that looks at the gender breakdown of commonly used wealth indexes.

This project sought to evaluate gender bias in AI applications in measuring wealth using household survey data. Data from the Ghana demographic and health survey, along with satellite imagery, were used to construct the module.

AI-based wealth estimation models have been shown to perform well in general across numerous studies, yet no work has explored their effectiveness at accurately accounting for conditions for subpopulations, such as women.

The findings are intended to inform more accurate poverty estimates that take gender into account when allocating resources, undertaking impact evaluations, and informing policy.

BACKGROUND

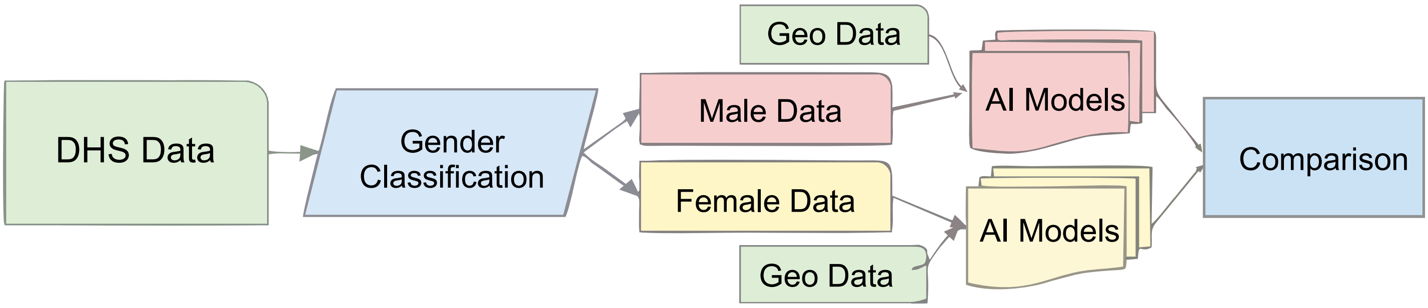
Over the past year, AidData, in partnership with CDD-Ghana, has worked to evaluate the potential of gender bias in wealth estimates generated using artificial intelligence (AI), geospatial data, and USAID's Demographic and Health Surveys (DHS) data. The project leverages AidData's expertise in AI, geospatial data, household surveys, and CDD-Ghana's knowledge of the local context and environments to produce a novel public good that will elevate equitability discussions surrounding the growing use of AI in development.

AidData was awarded funding for the project through USAID's Equitable AI Challenge. The Equitable AI Challenge - implemented through DAI's Digital Frontiers- was designed to fund approaches that will increase the accountability and transparency of AI systems used in global development contexts. The project builds upon AidData's broader research initiative on gender equity in development, as well as ongoing AI applications.

APPROACH

The DHS Wealth Index, an asset-based metric of household wealth, is one of the most widely used sources of training data for AI models which estimate wealth. AI-based wealth estimation models have been shown to perform well in general across numerous studies, yet no work has explored their effectiveness at accurately accounting for conditions for sub-populations, such as women. To explore variation in model accuracy for households led by women or men, we classify households surveyed in the 2014 Ghana DHS by gender and train separate models in order to compare them.

Since DHS assets are only recorded for the entire household, accounting for gender-specific conditions can be difficult. We classify households as male or female using varying methods, such as the gender of the head of household. Gender-specific AI models are trained using the Wealth Index as the dependent variable, and a range of geospatial data from satellite



We find that when using equal household counts, male models still outperform female models, suggesting gender is still meaningfully influencing model performance.

imagery and other sources as the independent variable. The geospatial data (including nighttime lights, population, land cover, and more), along with the code used to train models, are publicly available to support replication and future use.

FINDINGS

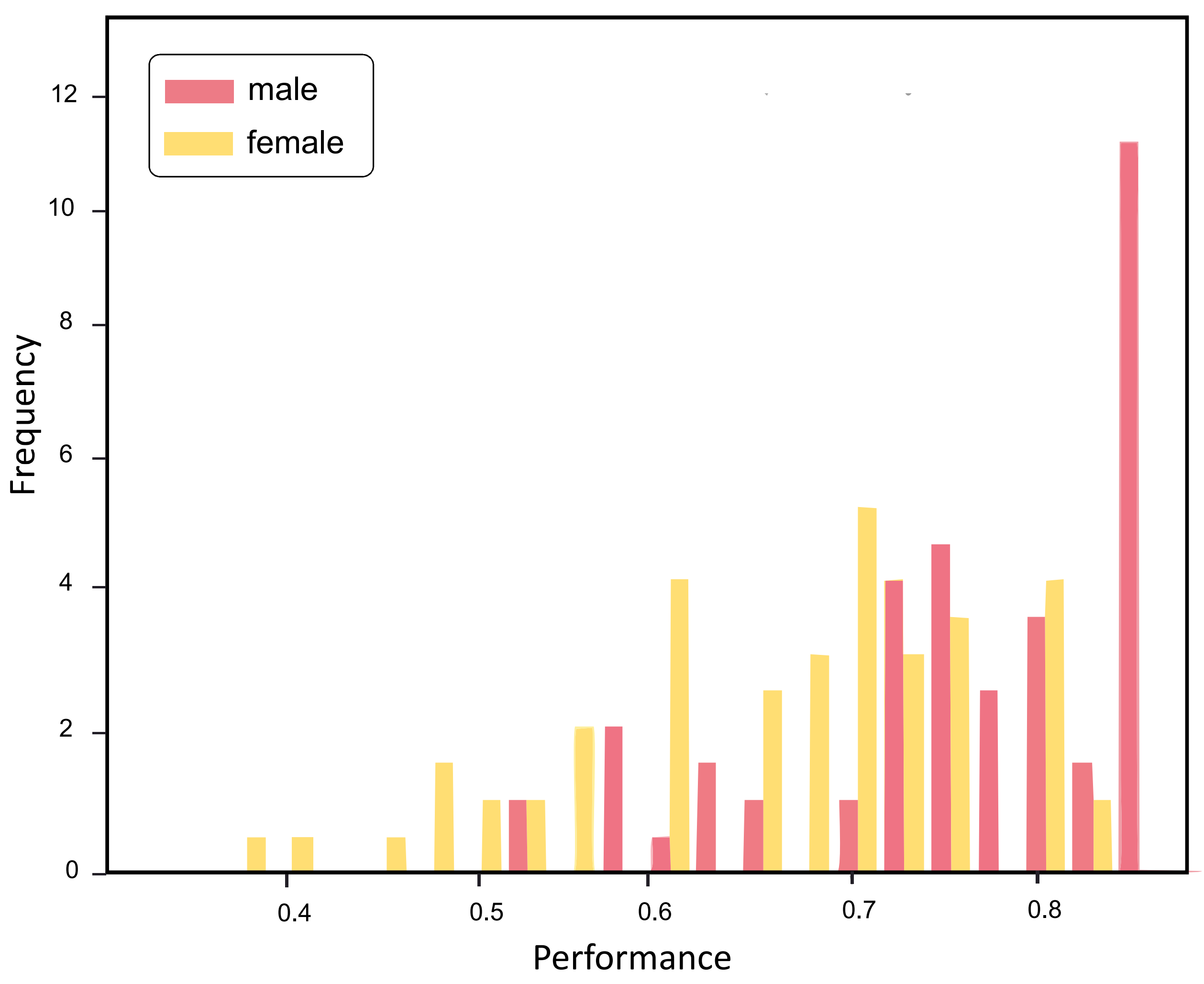
We test the performance of the AI models trained across a range of parameters and inputs that impact model behavior, using data for each gender. Across tests, models trained on male household data consistently outperform - or more accurately estimate wealth- models trained on female household data.

A critical factor is the differences in the number of male and female households used to train the models. As there are typically more male households, we run robustness tests in which we limit the number of households to be equal across genders. We find that when using equal household counts, male models still outperform female models, suggesting gender is still meaningfully influencing model performance.

A valuable feature of the type of AI models used is the ability to measure the importance of specific features (the geospatial variables) used to estimate wealth in the models. Male and female models had similar feature importance overall, and were most heavily influenced by a subset of features, including nighttime lights, urban area coverage, and population. Notable differences include male models being more influenced by rainfall amount, while female models were influenced by distance to cities.

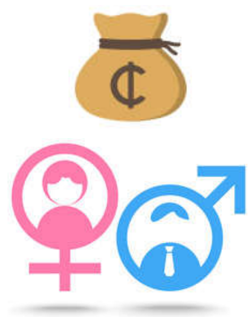
Additional efforts to understand the influence of gender on the DHS Wealth Index itself involved rebuilding the index using gender-specific data. We found that the gender-specific wealth index created for female-led households tended to classify poorer households as even poorer relative to the original DHS WI.

Comparison of Gender Model Performance



FUTURE DIRECTIONS

Our current research has indicated that AI models trained on female household data underperform models trained on male household data, yet there are many aspects left to explore. An important area for consideration is whether current household gender classification is appropriate, and, more broadly, whether future surveys can be improved to assess gender-specific wealth. Understanding what drives the differences between models trained on male and female data is also important- can other geospatial data features used in model training improve the performance of female models? Future research might also explore the possibility of utilizing wealth indices other than those produced by the DHS. The methods and code we have produced will hopefully provide an accessible approach for others to continue exploring these questions and others related to the role of gender in AI wealth estimation models.



ARE CURRENT HOUSEHOLD GENDER CLASSIFICATIONS APPROPRIATE?
CAN FUTURE SURVEYS BE IMPROVED TO ASSESS GENDER-SPECIFIC WEALTH?

The views expressed are solely the responsibility of the authors and do not necessarily reflect the views of the development partners.

Contact Information:

AidData - Rachel Sayers: rsayers@aiddata.org

CDD-Ghana - Mavis Zupork Dome: m.dome@cddgh.org