

The Global Program on GovTech and Public Sector Innovation supports the adoption of technology and use of data to enhance government effectiveness and efficiency, as well as increase citizen engagement, and improve governance. Learn more: www.worldbank.org/govtech

Challenge

Tax administrations globally face difficulties keeping up with modern economies and sophisticated tax evasion tactics. Traditional rule-based systems often fall short in detecting these tactics effectively. Machine learning (ML) models can improve audit case selection, but their training requires extensive prior data and collaboration, which is often hindered by privacy concerns and a lack of standardized reference datasets. Tax administrations often start from scratch without shared AI models or published performance reports. Testing various AI architectures

is time-consuming, and AI solutions have limited validation.

Solution

To address these issues, the World Bank's Gov-Tech Innovation Lab created a synthetic data generator that uses algorithms and statistical methods to mimic real-world tax operations data. This allows for the training of machine learning algorithms without compromising data privacy. A prototype ML algorithm was designed to select cases for tax audits, and it was initially validated using anonymized data from the Georgia Revenue Service, achieving a promising accuracy rate. This hybrid approach, which combines synthetic data for initial training and country-specific data for fine-tuning, shows significant potential for improving tax evasion detection and enhancing tax audit efficiency, boosting tax revenues, increasing risk perception, and reducing administrative costs.

Related resources:

Governance for Development Blog
How is the World Bank using Al
and Machine Learning for Better
Governance? ☐



