XINGU: EXPLAINING CRITICAL GEOSPATIAL PREDICTIONS IN WEAK SUPERVISION FOR CLIMATE FINANCE

David Dao*, **Johannes Rausch***, **Iveta Rott and Ce Zhang**
Department of Computer Science
ETH Zurich
Zurich, Switzerland
{david.dao,jrausch,rotti,ce.zhang}@inf.ethz.ch

ABSTRACT

Monitoring, Reporting, and Verification (MRV) play a crucial key role in the decision-making of climate investors, policymakers and conservationists. Remote sensing is commonly used for MRV but practical solutions are constrained by a lack of labels to train machine learning-based downstream tasks. Recent work leverages weak supervision to alleviate the problem of labeled data scarcity. However, the definition of weak supervision signals is limited by the existence of millions of possible heuristic-based feature generation rules. Furthermore these rules are often difficult to interpret for climate finance and underperform in critical data subsets. We propose Xingu, an interpretable MRV system to explain weak supervision rules using game-theoretic SHAP values for critical model predictions. Moreover Xingu enables domain experts to collectively design and share labeling functions, thus curating a reusable knowledge base for weak supervision signals.

1 INTRODUCTION

Humanity is facing an unprecedented climate crisis and decisive action is required to tackle global warming. Land use and land use change (LULUC) play a critical role in our climate, taking up about a quarter of annual anthropogenic emissions of greenhouse gases (GHGs) during 2007-2016 (IPCC (2019)). In addition to being a key driver of global warming, careless land use is also destroying valuable ecosystem services and is threatening the livelihood for local populations and a multitude of species.

Multiple climate finance instruments have been established to reward landowners for sustainable land use REDD+, Payment for Ecosystem Services and biodiversity banking (Díaz et al. (2015)). A key task of global climate finance is the Monitoring, Reporting and Verification (MRV) of global change in land use, forests and biodiversity (Dao et al. (2019a)). Aircraft- or satellite-based MRV is crucial for the decision making of stakeholders such as government agencies and private climate investors. Machine learning approaches have shown to be effective to automatically perform MRV-tasks, but are limited by a lack of task-specific labels. Due to this lack of available labeled data, recent work has thus been focusing instead on leveraging weak supervision to generate training data via labeling functions (Ratner et al. (2016), Dao et al. (2019b)).

However, weak supervision systems are facing a multitude of challenges that make decision making within climate finance difficult:

1. The enormous availability of geospatial data is restricting domain experts in building strong labeling rules. For instance, Google's Earth Engine data catalog provides 376 datasets (as of January 2020), each providing multiple possible bands. Facing millions of possible combinations and preprocessing rules, domain experts struggle to select and design the right weak label generators.
2. Being able to interpret model predictions is an important requirement for transparent and fair decision making in climate finance. However outcomes from an MRV system trained

*Authors have contributed equally.

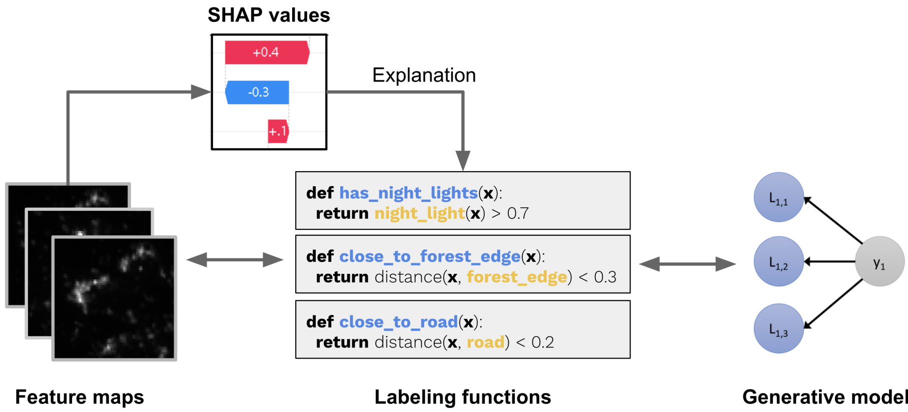

Figure 1: Local explanations of labeling functions. In data programming, labeling functions are used to estimate a generative model in order to predict probabilistic labels. Surprisingly, on the same time geospatial labeling functions can also be represented as ternary feature maps. Thus the role and relationship of labeling functions towards a prediction on labeled critical data subsets can be approximately explained using their corresponding SHAP values.

with weak supervision are difficult to explain. Especially the role of labeling functions in data programming predictions is not well understood.

3. Especially critical outcomes for climate finance (e.g selective logging or mining in LULUC) correspond to, often rare, data subsets or slices (Chen et al. (2019)). While MRV models that leverage weak supervision can achieve overall high accuracy, they may underperform on critical subsets due to lack of slice-specific training data.

4. Weak supervision approaches that are constrained to the image domain are often unable to incorporate highly informative location-specific, socio-economic or political data. For instance, weak supervision such as night lights (Jean et al. (2016)) or Doc2Vec representations extracted from geotagged Wikipedia articles (Sheehan et al. (2018)) can be used as features for socio-economic growth prediction.

We propose Xingu: An interpretable automated MRV system that addresses the aforementioned issues with respect to feature design and transparency issues in climate finance:

1. Xingu enables domain experts to design and share labeling functions for specific tasks. Understanding the role of labeling functions of a certain task then allows us to recommend these functions for novel but similar tasks.

2. Xingu incorporates SHAP (Lundberg & Lee (2017)), a method that allows measurement of feature-importance on a given dataset. As such, features can be scored and interpreted in terms of their usefulness. All labeling functions that are used in our system are defined as ternary maps (corresponding to *yes*, *no*, *abstain*). As such, we can obtain additional information about the importance of labeling functions by using them as feature inputs for tree-based models on critical data slices and explaining their predictions via their SHAP values. As we see in Section 2, this provides us with a novel way to investigate the role of labeling functions in critical prediction tasks.

3. Our system provides an interface that allows stakeholders to contribute small labeled samples of performance-critical regions. Subsequently, expert features can be analyzed in terms of their individual impact on the region samples.

4. Xingu enables domain experts to build features that incorporate political or socioeconomic information. For example, features that make use of *night lights* satellite images could be impactful in the context of training a deforestation classifier, as they are an indicator of human activity.

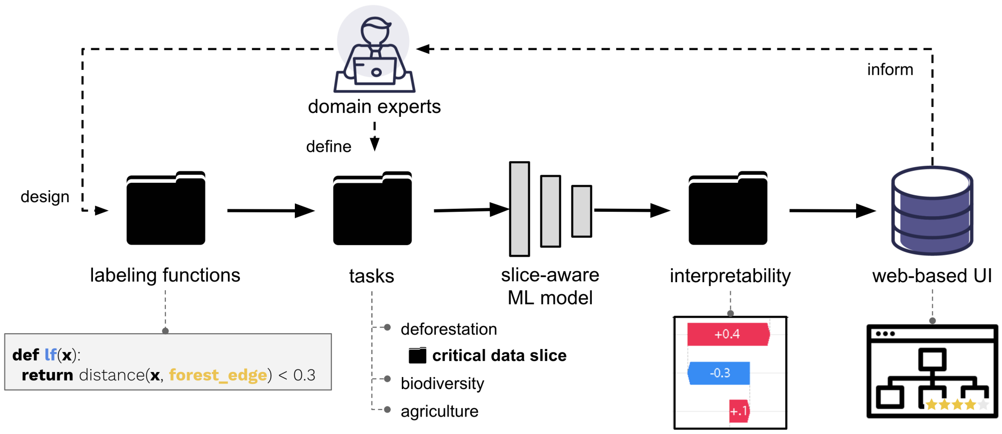

Figure 2: Xingu's workflow. Domain experts are encouraged to design labeling functions and submit task-specific critical data slices with the community. Using data programming, a slice-aware ML model is then trained and interpreted. Finally the results are displayed in an interactive web-based UI.

2 Understanding critical predictions with feature importance

Xingu applies leverages weakly labeled datasets by allowing users to define labeling functions instead of hand-labeling individual examples (Ratner et al. (2017)). Based on overlaps of multiple labeling functions for given data points, a generative model is trained in order to predict probabilistic labels. In contrast to other work that utilized weak supervision rules to generate noisy training labels, Xingu further aims to explain the role of individual weak supervision rules by using local SHAP value explanations on critical data slices. These explanations are used as a basis for an an interactive report that is provided to domain experts.

2.1 Interpreting geospatial labeling functions using SHAP

SHAP is a game-theoretic approach to explain the output of model using the Shapley value. SHAP values explain how much a feature pushes the model output from the base value (the average model output over the training dataset we passed) to the model output. In order to leverage SHAP for labeling functions, we explore the surprising note that every labeling function for geospatial input can also be represented as a tenary feature map. Finally we use this feature map representation of a labeling function as input to SHAP's tree explainer model to extract its corresponding SHAP values on the provided critical data subsets as shown in Figure 1.

2.2 Curating a web-based knowledge base

Xingu workflow as shown in Figure 2 enable domain experts to interactively submit their custom labeling functions and explore its explanations. Domain experts can also decide to share their labeling functions, processing scripts and a description with a public community. The community can then rate the quality of contributions and discuss them in a public forum. Domain experts can also decide to share their tasks and critical data slices (represented as a collection of GeoJSON) with the public. Xingu's goal is to enable domain experts to curate a knowledge base of good labeling functions and critical data slices over time that can be reused, either as templates for future geospatial analysis projects on related scenarios or as baselines for further explorations on the same scenario. Collectively curating labeling functions provides an opportunity to not only understand the relationship between these functions (e.g. exclusive, overlapping) and their impact, but also to recommend suitable labeling functions for novel tasks using task similarity (Zamir et al. (2018)).

3 Discussion

Xingu facilitates training of machine learning-based MRVs for various downstream tasks by leveraging domain expertise of critical data regions. Over the coming years, geospatial data is expected to grow rapidly, with more and more observation infrastructure in place (smaller satellites, more cost-efficient launching rockets). Better technology deployed as part of this growing infrastructure will deliver higher resolution images, as well as different non-image data streams. An interpretable and adaptable MRV system is therefore crucial to guide climate investors and policymakers in their decision-making.

Acknowledgments

The authors are thankful for the guidance and advise by the CONAF (Daniel Montaner, Cesar Mattar, Jose Antonio Prado). Part of this research has been developed as part of the OpenSurface platform and a real-world pilot in Chile, which was launched at the COP25 United Nation's Climate Summit. OpenSurface is funded by IDBLab and EIT Climate-KIC.

References

Vincent S. Chen, Sen Wu, Zhenzhen Weng, Alexander Ratner, and Christopher Ré. Slice-based learning: A programming model for residual learning in critical data slices, 2019.

David Dao, Catherine Cang, Clement Fung, Ming Zhang, Nick Pawlowski, Reuven Gonzales, Nick Beglinger, and Ce Zhang. GainForest: Scaling Climate Finance for Forest Conservation using Interpretable Machine Learning on Satellite Imagery. *ICML Climate Change AI workshop 2019*, 2019a.

David Dao, Johannes Rausch, and Ce Zhang. Geolabels: Towards efficient ecosystem monitoring using data programming on geospatial information. In *2019 NeurIPS Workshop on Tackling Climate Change with AI*, 2019b.

Sandra Díaz, Sebsebe Demissew, Julia Carabias, Carlos Joly, Mark Lonsdale, Neville Ash, Anne Larigauderie, Jay Ram Adhikari, Salvatore Arico, András Báldi, et al. The IPBES Conceptual Framework—connecting nature and people. *Current Opinion in Environmental Sustainability*, 14:1–16, 2015.

IPCC. 2019: Summary for policymakers. In P.R. Shukla, J. Skea, E. Calvo Buendia, V. Masson-Delmotte, H.-O. Pörtner, D. C. Roberts, P. Zhai, R. Slade, S. Connors, R. van Diemen, M. Ferrat, E. Haughey, S. Luz, S. Neogi, M. Pathak, J. Petzold, J. Portugal Pereira, P. Vyas, E. Huntley, K. Kissick, M. Belkacemi, and J. Malley (eds.), *Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems*, pp. 7–11. 2019.

Neal Jean, Marshall Burke, Michael Xie, W. Matthew Davis, David B. Lobell, and Stefano Ermon. Combining satellite imagery and machine learning to predict poverty. *Science*, 353(6301): 790–794, 2016. ISSN 0036-8075. doi: 10.1126/science.aaf7894. URL `https://science.sciencemag.org/content/353/6301/790`.

Scott M Lundberg and Su-In Lee. A unified approach to interpreting model predictions. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett (eds.), *Advances in Neural Information Processing Systems 30*, pp. 4765–4774. Curran Associates, Inc., 2017. URL `http://papers.nips.cc/paper/7062-a-unified-approach-to-interpreting-model-predictions.pdf`.

Alexander Ratner, Stephen H. Bach, Henry Ehrenberg, Jason Fries, Sen Wu, and Christopher Ré. Snorkel: Rapid Training Data Creation with Weak Supervision. *Proc. VLDB Endow.*, 11(3): 269–282, November 2017. ISSN 2150-8097. doi: 10.14778/3157794.3157797. URL `https://doi.org/10.14778/3157794.3157797`.

Alexander J Ratner, Christopher M De Sa, Sen Wu, Daniel Selsam, and Christopher Ré. Data programming: Creating large training sets, quickly. In *Advances in neural information processing systems*, pp. 3567–3575, 2016.

Evan Sheehan, Burak Uzkent, Chenlin Meng, Zhongyi Tang, Marshall Burke, David Lobell, and Stefano Ermon. Learning to interpret satellite images using wikipedia, 2018.

Amir Roshan Zamir, Alexander Sax, William B. Shen, Leonidas J. Guibas, Jitendra Malik, and Silvio Savarese. Taskonomy: Disentangling task transfer learning. *CoRR*, abs/1804.08328, 2018. URL `http://arxiv.org/abs/1804.08328`.