UC San Diego | SCRIPPS INSTITUTION OF OCEANOGRAPHY

BERT CLASSIFICATION OF PARIS AGREEMENT CLIMATE ACTION PLANS

Tom Corringham[1], Daniel Spokoyny[2],
Eric Xiao[3], Christopher Cha[3], Colin Lemarchand[3],
Mandeep Syal[3], Ethan Olson[3], Sasha Gershunov[1]

[1]Scripps Institution of Oceanography[2]
Carnegie Mellon University [3]University of California San Diego

Using NLP to extract policy information from climate documents

- The volume of text-based information on climate policy is increasing.
- NLP tools can distill information from text to inform climate policy.
- We seek to classify sentences in national climate action plans by topic, e.g., mitigation, adaptation, energy, land use, etc.
- Gold standard: use human-annotated sentences to train model. But human expert annotators are expensive.
- Our framework: exploit the header structure in the climate action plans to generate noisy topic labels.
- Train the BERT (Devlin et al., 2018) state-of-the-art NLP deep learning model using these noisy labels.
- Evaluate BERT performance against a heuristic bag-of-words classifier.
- Assess how well the BERT classifier agrees with untrained and trained human annotators.

Paris Agreement Nationally Determined Contributions (NDCs)

ISLAMIC REPUBLIC OF AFGHANISTAN

Intended Nationally Determined Contribution

Submission to the United Nations Framework Convention on Climate Change

21 September 2015

*****

The Islamic Republic of Afghanistan hereby communicates its Intended Nationally Determined Contribution (INDC) and information to facilitate understanding of the contribution.

Executive Summary	
Base Year:	2005
Target Years:	2020 to 2030
Contribution Type:	Conditional
Sectors:	Energy, natural resource management, agriculture, waste management and mining
Gases Covered:	Carbon dioxide (CO_2), methane (CH_4), and nitrous oxide (N_2O)
Target:	There will be a 13.6% reduction in GHG emissions by 2030 compared to a business as usual (BAU) 2030 scenario, conditional on external support.
Financial Needs:	Total: USD 17.405 billion • Adaptation: USD 10.785 billion • Mitigation: USD 6.62 billion (2020-2030)

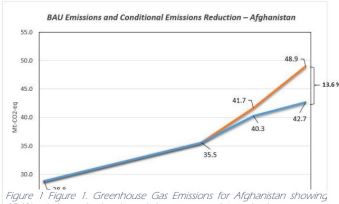

Figure 1 Figure 1. Greenhouse Gas Emissions for Afghanistan showing 13.6% relative reduction in emissions compared to a business as usual scenario for the year 2030

1. Afghanistan's National Circumstances and Commitment to Climate Change

Introduction

Afghanistan has extensive development and climate adaptation needs and, currently, low levels of greenhouse gas (GHG) emissions. Afghanistan remains one of the poorest countries in the world today, with an estimated population of 28.6 million (2015)[1] and a per capita GDP of USD 660.[2]

Afghanistan is highly prone to natural disasters throughout its 34 provinces.[3] As a result of climate change, it is anticipated that the incidence of extreme weather events, including heat waves, floods, and droughts will likely increase, as will climate change-linked disasters such as glacial lake outflows. The majority of Afghanistan's population relies directly or indirectly on the available natural resources for their livelihoods so with these climatic changes the foundation of the country's economy, stability, and food security is under threat.

Despite these challenges, Afghanistan can remain a low emission economy while developing rapidly if, under the Paris Climate Change Agreement, extensive financial and other resources are made available to allow Afghanistan to successfully develop and implement Low Emission Development Strategies (LEDS) and Highly Effective Adaptation and Development Strategies (HEADS).

Appropriate support in the form of finance, capacity building, technology and legal assistance is needed for Afghanistan to make substantial progress on social and economic fronts while maintaining low per capita GHG emission levels.

2. Climate Change Adaptation

Near- and Long-term Adaptation Visions, Goals and Targets

Afghanistan's vision for addressing the adverse impacts of climate change through adaptation aims to protect the country and its population by enhancing adaptive capacity and resilience, effectively respond to the vulnerabilities of critical sectors, and efficiently mainstream climate change considerations into national development policies, strategies, and plans. In order to achieve this vision, a national strategy for climate change adaptation must include community level vulnerabilities and build up their adaptive capacities by investment in short- and long-term initiatives. Short-term action plans formed part of the 2009 NAPA, while the NAP will implement both short- and long-term priorities These priorities include, but are not limited to:

1. Reducing vulnerability of the country and its population through enhancement of adaptive capacity and resilience, and deployment of disaster risk reduction approaches
2. Integrating climate change consideration into the national planning processes
3. Promoting economic development and sustainable rural livelihoods through sustainable management of environmental resources and increase access to modern forms of efficient and sustainable energy services
4. Improvement of technical capacity in governmental institutions
5. Adaptive and integrated land and water management
6. Improving access by rural communities and farmers to water to support food security, reduce poverty and improve agricultural productions
7. Raising awareness for people of Afghanistan on climate change impacts and adaptation measures

3. Climate Change Mitigation

GHG Emissions and Mitigation Measures

Afghanistan has very low relative per capita GHG emissions. While 1990 emissions were at 0.2 metric tons CO_2 per capita, data indicates that per capita emissions were around 0.3 for 2010, making Afghanistan one of the lowest GHG emitters globally.[7] However, the country is on a growth path, which is expected to strengthen over the coming years, meaning GHG emissions are likely to increase. It is important that support be provided to Afghanistan to develop LEDS to minimize the increase in its GHG emissions.

Afghanistan's overall GHG emission figures (Table 1) demonstrate that the most important sources of CO_2 emissions are from the "Land-Use Change and Forestry" and "Energy" sectors.[8] In terms of CH_4 and N_2O emissions, the agriculture sector is the major contributor.[9] It is therefore important that greater attention should be focused on GHG mitigation measures within these sectors. Afghanistan has only started to access UNFCCC technology transfer opportunities. It needs to build national capacity to navigate opportunities and play a more critical role in international negotiations.
The following mitigation options are designed to enable Afghanistan to make a mitigation contribution which is condition on support needs for financial and technical support being met.

Climate Mitigation Gaps and Barriers and Support Needs USD 662 Million/Year from 2020		
Sector	**Technology and Capacity Building Needs**	**Finance Needs (USD)**
Energy Efficiency in Buildings and in Transport Sector	Carbon finance and project development skills. Information on available technologies, measures, and financing skills. Traditional customs and administered pricing. Building codes, and standards on appliances and equipment. Clean cooking, heating and power projects.	100 million/Year
Energy	Human and institutional capacity for adoption of cleaner technology. Capital markets that encourage investment in decentralized systems. Information and intellectual property rights for mitigation technologies. Renewable energy, entry costs support, access to capital, and subsidies. Environmental compliance standards (emission and indoor).	188 Million/Year
Waste Management	Landfill management, decentralised wastewater treatment. Climate Project development skills.	74 Million/Year
Forest and Rangelands	Carbon sequestration on forest/rangelands, and forest carbon skills. Funding institutional capacity to monitor and verify projects. Better spatial planning for community and production agriculture. Reduce rural peoples' dependence on fuel for cooking and heating.	100 Million/Year
Industry and Mining	Cleaner coal mining, leave-it-in-the-ground approaches, combustion, and transportation of minerals. Hydrocarbon fields management. Technical industrial capacity to link basic industry and mining private and public sector with climate sector experts.	100 Million/Year
Agriculture and Livestock	National herd, reduction in fuel used, or cleaner fuel technologies. South-south collaboration on low-carbon agriculture, study tours. Funding for R&D activities. Improved national dataset on agriculture, food security data.	100 Million/Year

Words in HTML headers mapped to topics by climate experts

1. Extract all words from headers and tables. Words from top row or left column extracted manually from tables.
2. Create bag-of-words mapping between words found in headers and topic labels. e.g., "land_use" ← "afolu | fires | forestry | forests | land | lulc | lulucf"
3. Link sentences from HTML paragraphs, lists and tables to topic labels keeping the most deeply nested label. In case of multiple labels choose the label with lower frequency.

Table 1. Frequency of weak labels over NDC sentences.

Weak Label	Frequency (%)
No Label	16.3
Adaptation	15.0
Agriculture	4.7
Economic	4.5
Energy	5.0
Environment	3.0
Equity	7.1
Industry	2.0
Land Use	3.4
Mitigation	16.0
Strategy	21.7
Waste	1.2

BERT model trained and evaluated on 25,500 sentences

- The weakly labeled sentences were split into 80% training, 10% validation, and 10% test sets.
- The BERT model (Devlin et al., 2018) was trained and tested against the data.
- Training continued until validation loss increased for three consecutive epochs; the model with the lowest validation loss was selected.
- This model was then applied to the hold-out test set of sentences to evaluate model skill (Table 2).

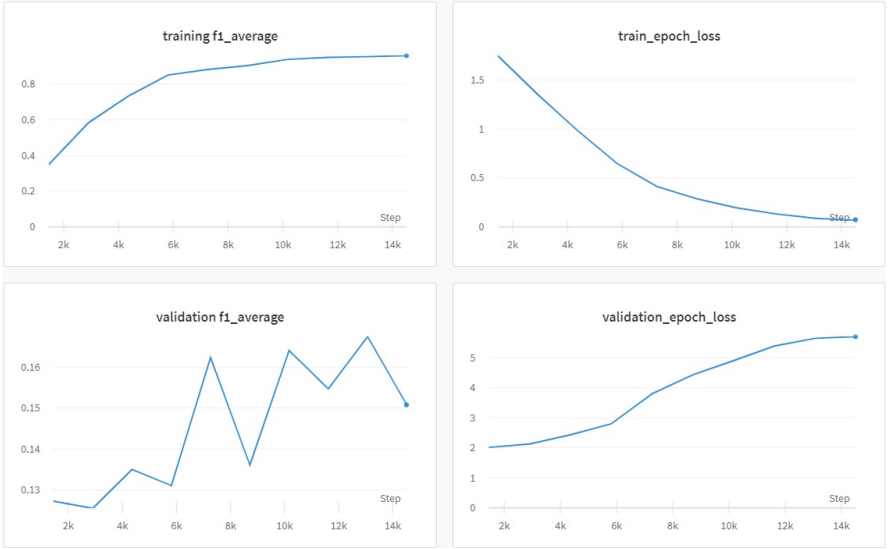

Table 2. BERT training skill scores.

Data	Accuracy	Precision	Recall	F_1
Train	0.907	0.692	0.673	0.669
Validate	0.839	0.450	0.436	0.429
Test	0.847	0.417	0.406	0.397

Evaluation of BERT prediction of weak labels relative to benchmarks

- We compared **$\text{BERT}_{\text{BASE}}$** and **SciBERT** transformer models to simple heuristics and human annotators (Table 3).
- **Random** randomly assigned labels to sentences with uniform distribution.
- **Majority** assigned the most common Strategy label to all sentences.
- The **Contains** heuristic assigned labels to sentences using the same bag-of-words mapping used to generate the HTML header labels.

Table 3. Model performance.

CLASSIFIER	ACCURACY	PRECISION	RECALL	F_1
RANDOM	0.813	0.117	0.081	0.089
MAJORITY	0.757	0.041	0.203	0.069
CONTAINS	0.829	0.229	0.170	0.171
SCIBERT	0.843	0.398	0.379	0.362
BERT	0.847	0.417	0.406	0.397
HUMAN*	0.867	0.281	0.250	0.251

- Two human annotators, a **Student** and a climate policy **Expert**, annotated 600 sentences from the test set by hand.
- The classifiers were all evaluated based on their ability to predict the weak HTML header-based labels.

Evaluation of BERT prediction of weak labels relative to benchmarks

Table 3. Model performance.

CLASSIFIER	ACCURACY	PRECISION	RECALL	F_1
RANDOM	0.813	0.117	0.081	0.089
MAJORITY	0.757	0.041	0.203	0.069
CONTAINS	0.829	0.229	0.170	0.171
SCIBERT	0.843	0.398	0.379	0.362
BERT	0.847	0.417	0.406	0.397
HUMAN*	0.867	0.281	0.250	0.251

Comparison of BERT and the Contains heuristic to human annotators

Table 4. Comparison of Contains and BERT to human annotators.

Weighted F_1		Reference Label		
		Human	Student	Expert
Classifier	Contains	0.350	0.399	0.302
	BERT	0.301	0.284	0.317
	Student			0.472

- **Contains** outperformed **BERT** in matching the **Student** sentence classifications (F_1 of 0.40 and 0.28, respectively).
- **BERT** outperformed **Contains** in matching the **Expert** classifications (F_1 of 0.32 and 0.30, respectively).

- **Student** may have relied more heavily on subject words while **Expert** may have integrated more context.
- That **BERT** outperforms **Contains** in matching **Expert** annotation is promising, despite low skill scores.
- The F_1 of 0.47 between **Student** and **Expert** classifications indicates a high level of ambiguity in sentence topics.
- In future work we will explore multi-label classification and larger sets of (possibly hierarchical) subject topics.

Comparison of BERT and the Contains heuristic to human annotators

Table 4. Comparison of Contains and BERT to human annotators.

WEIGHTED F_1		REFERENCE LABEL		
		HUMAN	STUDENT	EXPERT
CLASSIFIER	CONTAINS	0.350	0.399	0.302
	BERT	0.301	0.284	0.317
	STUDENT			0.472

Comparison of BERT and the Contains heuristic to human annotators

Table 4. Comparison of Contains and BERT to human annotators.

WEIGHTED F_1		REFERENCE LABEL		
		HUMAN	STUDENT	EXPERT
CLASSIFIER	CONTAINS	0.350	0.399	0.302
	BERT	0.301	0.284	0.317
	STUDENT			0.472

Comparison of BERT and the Contains heuristic to human annotators

Table 4. Comparison of Contains and BERT to human annotators.

WEIGHTED F_1		REFERENCE LABEL		
		HUMAN	STUDENT	EXPERT
CLASSIFIER	CONTAINS	0.350	0.399	0.302
	BERT	0.301	0.284	0.317
	STUDENT			0.472

Error analysis

Table 5. Error analysis.

SENTENCE	LABEL	CONTAINS	BERT	STUDENT	EXPERT
It is envisaged that **emission** reduction will be achieved through the **mitigation** actions in the sectors.	MITIGATION	MITIGATION	MITIGATION	MITIGATION	MITIGATION
The Steering Committee is the supreme body for **decision making** and sectoral **implementation**.	STRATEGY	MITIGATION	STRATEGY	STRATEGY	STRATEGY
The **mitigation** actions that enhance **afforestation** are projected to result in the **sequestration** of 1 mtCO2e annually.	LAND USE	LAND USE	AGRICULTURE	MITIGATION	LAND USE
In the absence of project activity, **fossil fuels** could be burned in **power plants** that are connected to the grid.	STRATEGY	EQUITY	ENERGY	INDUSTRY	ENERGY
Due to the outbreak of the Ebola Virus the development gains made after a 10-year civil war were rudely reversed.	ENVIRONMENT	NO LABEL	MITIGATION	NO LABEL	NO LABEL

Future research

Natural Language Processing

- Multiple topic labels may provide improvements over the single topics.
- Multiple weak HTML header-derived labels could be used in training and BERT output layer weights could be evaluated against heuristic and human-labeled multi-labels using Jaccard similarity or other metrics.
- Here, topics and keywords were selected manually by climate experts. This task could be automated.

Climate Policy Evaluation

- Information from NDCs could be compared with country-level characteristics to investigate a number of policy questions.
- e.g., How do measures of GDP, energy intensity, or climate vulnerability compare to NLP-derived NDC characteristics?
- Are BERT-derived document characteristics more informative than simple word frequency metrics?

“Act as if the house is on fire because it is” – Greta Thunberg

“Climate change is a global problem” – Pope Francis

“Climate change knows no borders” – Angela Merkel, German Chancellor

“There is no planet B” – Emmanuel Macron, French President

“I’m not denying climate change…” – Donald Trump, former U.S. President

“Number one issue facing humanity” – Joe Biden, U.S. President