

TCFD-NLP: Assessing alignment of climate disclosures using NLP for the financial markets

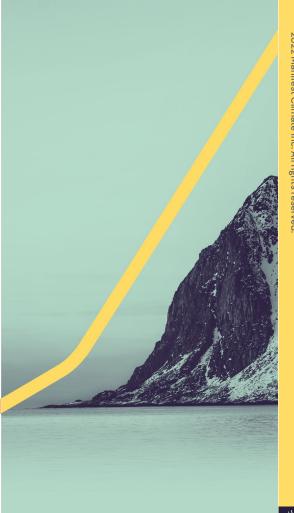
Rylen Sampson Aysha Cotterill Quoc Tien Au

Manifest Climate

Climate risk planning startup

We support organizations in understanding, managing, and communicating climate risks and opportunities

Software product supported by climate expertise



Climate-related financial disclosures

Increased investor and regulatory pressure has led to climate reporting frameworks including the Task Force on Climate-Related Financial disclosures (TCFD)



TCFD has 11 recommendations

| Governance | Strategy | Risk Management | Metrics and Targets |
|-----------------------------|--|---|--------------------------|
| Board supervision | Climate risks and opportunities | Processes for identifying and assessing climate risks | Climate metrics |
| Management responsibilities | Impacts from climate risks and opportunities | Processes for managing climate risks | Greenhouse gas emissions |
| | Climate scenarios | Integration of climate risks management | Climate targets |



Climate Disclosure Reviews

Tack Climate Change Outlook 202 Our Position and Policy on Climate Change and **Our Commitments to Climate Action** At Teck, we believe that climate change is a key global risk that is directly influenced by human activity and that requires decisive global action. Failure to act will expose the world to climate change impacts that will be costly for global ecosystems and for society as a whole. We believe we have a responsibility to help address this global challenge by reducing emissions at our operations and in our value chain, and by sustainably producing the metals, minerals and energy that are essential for building the technologies and infrastructure needed to transition to a low-carbon economy. related risks and advancing opportunities in our business strategies. Our Board of Directors and senior management the resilience of their communities and local ecosystems to are involved in assessing climate-related risks and opportunities the potential physical impacts of climate change to enable Tack to plan for these business and market forces Work with governments and other stakeholders to and to maintain resilience. We recognize that timely and accelerate the global response to climate change and change are of importance to Teck and our communities along with the positions of the organizations of which we As a company committed to climate action, we will: Provide timely and transparent of Integrate consideration of climate-related risks and related risks and opportunities opportunities into our strategic planning processes in addition to strong sustainability performance, timely and Reduce our operational greenhouse gas emissions in line transparent disclosures are also of importance to Teck and ou communities of interest. With respect to disclosures pertaining Set emissions reductions targets that provide transparency to climate change, the Task Force on Climate-related Finance as to how we will deliver reductions consistent with limiting recommendations in June 2017 for how companies can global warming to 15°C improve climate-related disclosures. We support the desire ecommendations and we signed on as a supporter of the Adapt to the potential physical impacts of climate chance

Annual Reports, Proxy Statements, ESG reports, etc

Disclosure review

Climate content - 11 TCFD categories

"Our Board of Directors and senior management are involved in assessing climate-related...." - **Board Supervision**

"In 2021, RBC EL used a tool called Transition Check to assess the transition risk impacts to a substantial portion of its wholesale loan exposures in the energy...." – **Climate Scenarios**

Management responsibility

Climate risks/opportunities

Integration of climate risk management

Climate metrics

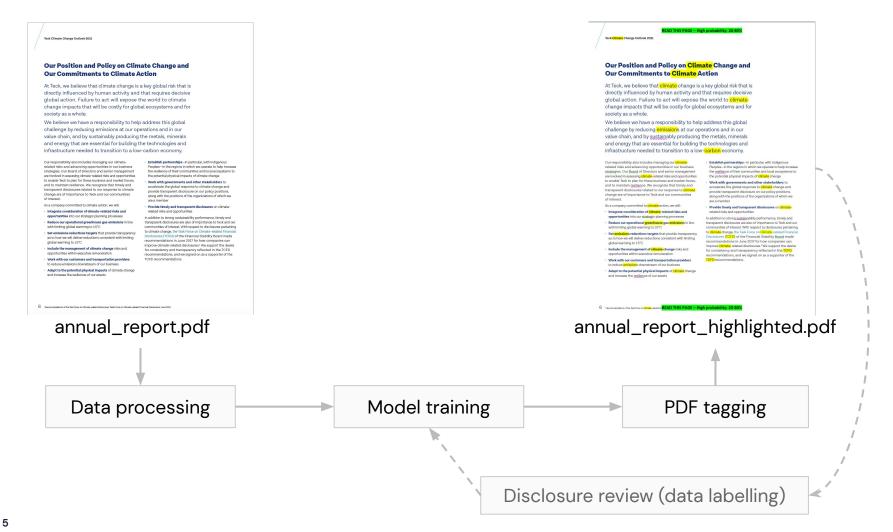
Climate targets

GHG emissions

•••

Non-climate content

| 1. | | | |
|----|------|------|--|
| | | | |
| 3. | | | |
| | | | |





Climate impact

Problem

Assessing TCFD-alignment is time-consuming (climate content is spread across many reports, often 100s of pages long)

Solution

Build NLP models to classify report text into categories

Impact

Enables large-scale analysis of corporate climate disclosure across sectors, geographies etc.

Helps companies to uncover gaps in how they are managing and communicating climate risks/opportunities

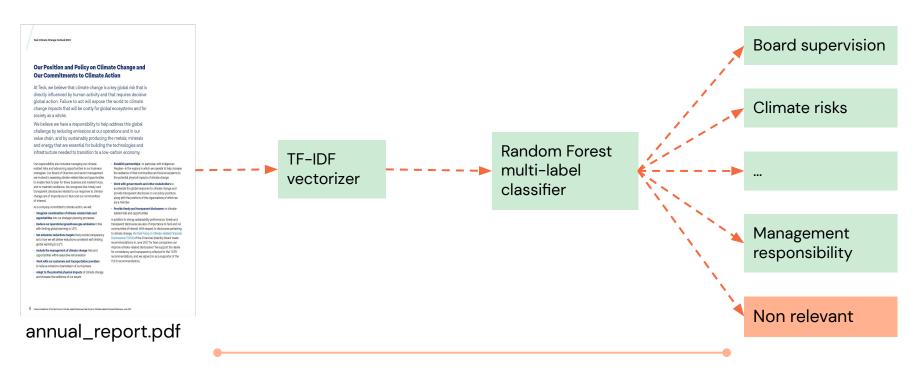


> 1000 reports labelled across industries

2 million sentences labelled with TCFD recommendations and non-relevancy



1. Baseline

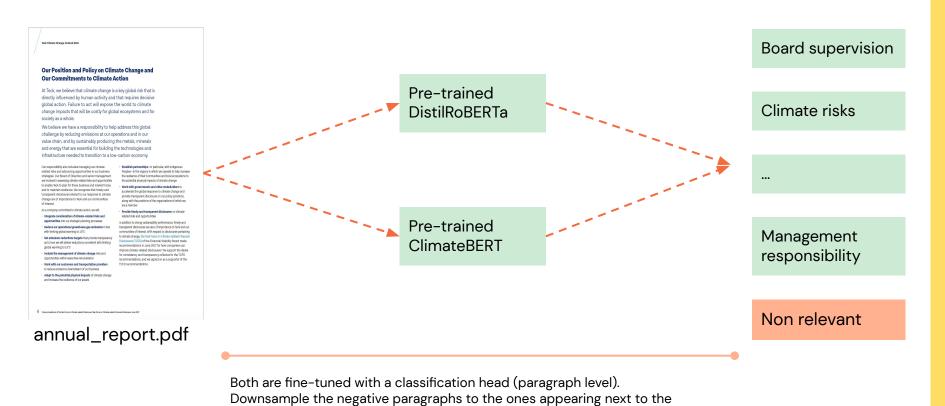


TF-IDF is trained on pages that contain TCFD-relevant information. RF is trained with a ratio of 2 negative pages for each positive page.

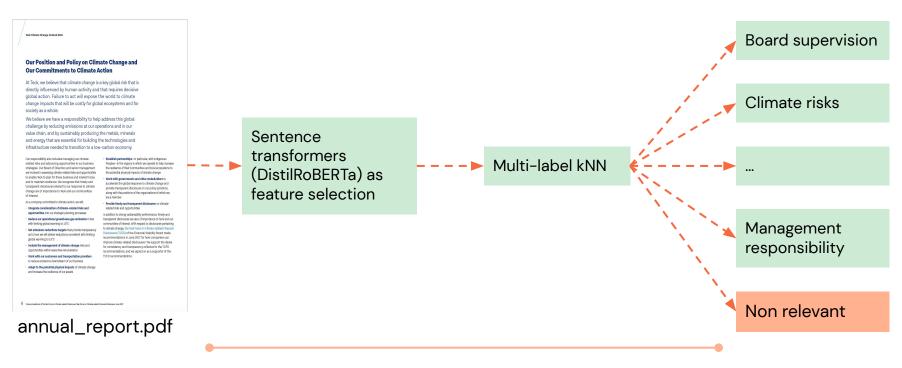


2. Fine-tuning language models

positive ones.



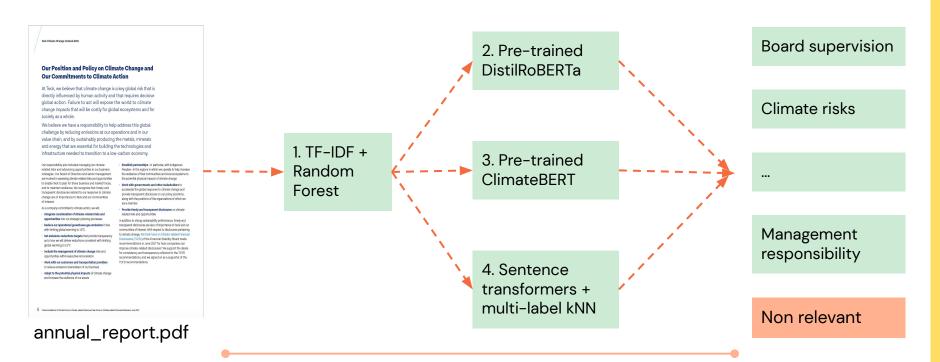
3. Sentence-transformers and multi-label kNN



The multi-label kNN uses the cosine distance.



4. Stacked models



The 3 models are only applied to the pages predicted as TCFD-relevant by the baseline.



Preliminary results

Table 1: ROC-AUC scores at the file level for every TCFD recommendation.

| File level ROC-AUC | TF-IDF (1) | kNN (2) | FT ClimateBERT (3) | FT distilRoBERTa (4) | (1) + (2) | (1) + (3) | (1) + (4) |
|-----------------------|------------|---------|--------------------|----------------------|-----------|-----------|-----------|
| Average AUC | 0.867 | 0.847 | 0.852 | 0.819 | 0.884 | 0.865 | 0.857 |
| Governance A | 0.891 | 0.854 | 0.865 | 0.854 | 0.91 | 0.887 | 0.875 |
| Governance B | 0.835 | 0.812 | 0.809 | 0.795 | 0.832 | 0.823 | 0.823 |
| Strategy A | 0.844 | 0.875 | 0.834 | 0.766 | 0.888 | 0.841 | 0.838 |
| Strategy B | 0.877 | 0.874 | 0.853 | 0.824 | 0.896 | 0.864 | 0.852 |
| Strategy C | 0.881 | 0.845 | 0.856 | 0.805 | 0.866 | 0.86 | 0.854 |
| Risk Management A | 0.874 | 0.919 | 0.89 | 0.845 | 0.935 | 0.893 | 0.874 |
| Risk Management B | 0.839 | 0.882 | 0.871 | 0.836 | 0.882 | 0.88 | 0.862 |
| Risk Management C | 0.864 | 0.941 | 0.862 | 0.836 | 0.931 | 0.892 | 0.882 |
| Metrics and Targets A | 0.859 | 0.758 | 0.869 | 0.79 | 0.895 | 0.888 | 0.876 |
| Metrics and Targets B | 0.932 | 0.8 | 0.87 | 0.864 | 0.875 | 0.877 | 0.882 |
| Metrics and Targets C | 0.843 | 0.773 | 0.795 | 0.796 | 0.814 | 0.816 | 0.813 |

Not enough data on each recommendation to fine-tune language models.

Fine-tuned ClimateBERT offers a boost in performance.

A mix of keyword and embedding based methods seems to be the best.



Next steps

Fine-tune language models

Given the results using ClimateBERT, we want to fine-tune the sentence transformers model with our data

Better combination of models

Experiment with different model stacking, e.g. per recommendation or per industry

Introduce domain knowledge

Find a way to augment the keyword-based model with more domain expertise rules

