**ORIGINAL ARTICLE**

# Machine learning for research on climate change adaptation policy integration: an exploratory UK case study

Robbert Biesbroek[1] (iD) · Shashi Badloe[2] · Ioannis N. Athanasiadis[3] (iD)

## Abstract
Understanding how climate change adaptation is integrated into existing policy sectors and organizations is critical to ensure timely and effective climate actions across multiple levels and scales. Studying climate change adaptation policy has become increasingly difficult, particularly given the increasing volume of potentially relevant data available, the validity of existing methods handling large volumes of data, and comprehensiveness of assessing processes of integration across all sectors and public sector organizations over time. This article explores the use of machine learning to assist researchers when conducting adaptation policy research using text as data. We briefly introduce machine learning for text analysis, present the steps of training and testing a neural network model to classify policy texts using data from the UK, and demonstrate its usefulness with quantitative and qualitative illustrations. We conclude the article by reflecting on the merits and pitfalls of using machine learning in our case study and in general for researching climate change adaptation policy.

**Keywords** Machine learning · Quantitative text analysis · Climate change adaptation · Policy and decision making · Mainstreaming · Artificial intelligence

## Introduction

Studies from across the globe suggest that countries, regions, and cities are increasingly developing dedicated climate change adaptation policies, strategies, and measures to adapt to current and projected climate change impacts. Examples include raising awareness of climate risks, developing novel financial schemes to increase resilience, changing existing legislation, and implementing 'hard' and 'soft' adaptation measures on the ground (Bauer et al. 2012; Clar and Steurer 2019.; EEA 2014; Henstra 2017; Lesnikowski et al. 2016, Lesnikowski et al. 2015;

Uittenbroek et al. 2019). Consequently, it has become increasingly clear that the success of adaptation actions is influenced by the ability of governments to integrate or 'mainstream' a focus on climate change across relevant sectors, domains, and levels (Runhaar et al. 2018). Mainstreaming adaptation—or adaptation policy integration—refers to the process whereby climate change concerns become an integral part of the structural dimensions of sectoral public bureaucracies (e.g. changing sectoral policies) and influence the ways in which public governance actors perceive the problem and consider climate change in their day-to-day activities (Biesbroek and Candel 2020). Adaptation to climate change impacts is not a stand-alone policy target but rather an issue that needs to be considered across all relevant sectors and levels (Wellstead and Stedman 2014). In general, adaptation policy integration aims to reduce trade-offs across sectors and to promote synergies; reduce under- and overreaction by departments, organizations, or ministries in response to climate change impacts; prevent inefficient investments of (scarce) resources; and promote coherence and consistency in implementing actions on the ground (Candel and Biesbroek 2016, 2018; Cejudo and Michel 2017; Lenschow et al. 2018; Maor et al. 2017; Tosun and Lang 2017).

✉ Robbert Biesbroek
  robbert.biesbroek@wur.nl

[1] Public Administration and Policy group, Wageningen University & Research, Wageningen, Netherlands

[2] Bioinformatics group, Wageningen University & Research, Wageningen, Netherlands

[3] Geo-Information Science and Remote Sensing Laboratory, Wageningen University & Research, Wageningen, Netherlands
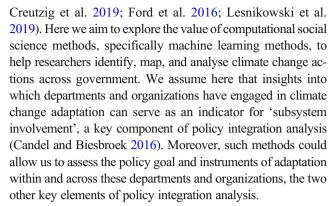
Whereas the need for integration has been increasingly recognized in policy practice and regarded in public policy theory as the holy grail for governing complex and cross-cutting policy problems (Peters 2015), empirical studies of climate policy integration, particularly focussing on adaptation, have been limited in scope and few in number. There are several reasons for this.

First, climate policy integration studies often focus on a limited number of departments and organizations, which are often the 'usual suspect' departments and organizations in adaptation: environment, water management, agriculture, and/or land use planning (Bauer and Steurer 2014; Runhaar et al. 2018; Uittenbroek et al. 2019). However, this represents a limited view of departments and organizations that, from a climate impacts perspective, should consider climate change, with sectors such as tourism, ICT, and transport often being ignored. Focussing on the usual suspect departments makes sense empirically and given resource constraints because there is a high(er) probability, and there is some evidence of adaptation policy integration taking place in these departments, but ignoring a large portion of relevant sectors and departments can become problematic as it does not show the full picture of how government is planning for and taking action on climate change adaptation.

Second, policy integration studies are often based on policy documents or evidence from interviews as empirical data (Tosun and Lang 2017). However, these are often limited in number and cover a short period of time, simply because it is too time-consuming and, thus, too costly to assess all policy documents ever produced by governments or to interview all relevant actors (Biesbroek et al. 2018). Related to this is the challenge of coding the findings in a systematic and transparent way, particularly when the number of documents significantly increases, thereby reducing (inter)coding reliability (Krippendorff 2018).

Third, it is not easy to identify climate change adaptation from policy texts. Although authoritative definitions of adaptation such as those provided by the IPCC (2014) exist, there are many alternative interpretations that differ across countries and regions, most noticeably between the climate change centric and social vulnerability centric interpretations (Dupuis and Knoepfel 2013). Consequently, most scholars have followed the strategy of including adaptation only when it is explicitly stated as such in text or referred to by interviewee respondents. This does not necessarily do justice to alternative interpretations of climate change adaptation which are likely to differ across countries. Contextual understanding of adaptation is therefore important to track how policy integration evolves and becomes integrated into existing bureaucracies.

Recent climate research and policy studies have made use of the rapid advancements in tools and methods developed and used in computational social sciences, particularly when it comes to 'text-as-data' methods (Biesbroek et al. 2018;

Creutzig et al. 2019; Ford et al. 2016; Lesnikowski et al. 2019). Here we aim to explore the value of computational social science methods, specifically machine learning methods, to help researchers identify, map, and analyse climate change actions across government. We assume here that insights into which departments and organizations have engaged in climate change adaptation can serve as an indicator for 'subsystem involvement', a key component of policy integration analysis (Candel and Biesbroek 2016). Moreover, such methods could allow us to assess the policy goal and instruments of adaptation within and across these departments and organizations, the two other key elements of policy integration analysis.

This article is structured as follows: In the next section, we will briefly introduce machine learning and document classification methods, highlighting its main assumptions, strengths, and weaknesses for adaptation policy research. In 'Creating and training the model' section, we elaborate our methodological design, including our justification for focussing on the UK. 'Evaluating the model performance' section presents the model performance and its evaluation. 'Using the algorithm to explore new UK policy documents' section demonstrates the model output and illustrates, using qualitative and quantitative examples, the value of the model. The article ends with a reflection on the potential of this approach and an outlook of next steps.

## Machine learning for adaptation policy research: a brief introduction

The interest in using computational social science methods in general and machine learning (ML) in particular has received increased attention in policy studies as it creates the possibilities for analysing large volumes of texts, images, sounds, and other types of data, allowing us to test existing theories and explore new ones (Anastasopoulos and Whitford 2019; Grimmer 2015; Lazer et al. 2009). But despite the growing popularity of ML, it has struggled to penetrate social science research on climate change with a few exceptions, for example, to explore discursive counter-movement networks in the USA (Farrell 2016), analyse social representations of adaptation (Lynam 2016), tracking donor funding (Donner et al. 2016), and analyse belief and sentiment using social media (Cody et al. 2015). Given its relative newness to climate change adaptation policy research, we highlight some key assumptions of ML. Since there are many useful and more elaborate resources available that provide a more in-depth discussion of ML (e.g. Hobson et al. 2019; Zizka et al. 2020), we merely highlight some key principles here that help to understand our model choice.

ML generally refers to a group of data-driven methods that combine algorithms and tools from computer science and statistics to 'learn' from data. Central to ML is the processes
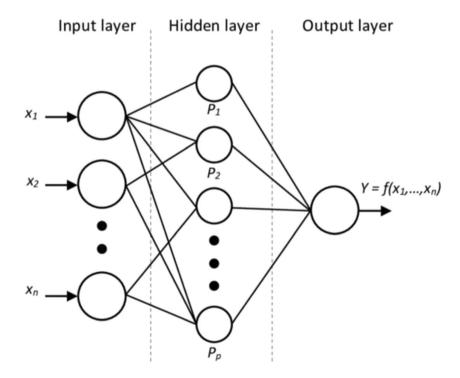
whereby the algorithm learns some patterns of a dataset (i.e. the 'training' dataset) and subsequently tests its performance on another dataset (i.e. the 'test' dataset). ML methods can adjust and optimize themselves based on previous data to perform better when confronted with new data. This process of learning from data is typically organized according to the kind of feedback provided to the algorithm (Russell and Norvig 2009). Supervised learning refers to the learning process whereby a labelled dataset is used, providing an answer key that the algorithm needs to learn reproducing. Such answer key usually consists of manually coded labels. Supervised learning is particularly useful for classification and regression problems. Classification problems are those where we want the algorithm to assign model inputs into one category or another, of example, whether a certain text fragment is predominantly 'mitigation' or 'adaptation' focussed. Regression problems are those where the model predicts a continuous valued output. Supervised learning only works if a clean, labelled dataset is available to train and test the algorithm, but existing labelled datasets are limited, and labelling data is time and resource intensive. In unsupervised learning, an algorithm is provided with a dataset without an answer key. Here the algorithm tries to structure the data to extract interesting patterns or associations. This type of learning model is particularly useful for clustering problems where the model aims to group unlabelled data. Computational social science methods and particularly machine learning for social science research on climate change predominantly employ unsupervised learning methods, as social network analysis, topic modelling, and sentiment modelling, as labelled datasets

are extremely scarce in social science research. Furthermore, the labelling process can be quite subjective in social sciences, in contrast with other domains where machine learning is thriving (Zizka et al 2020; Martin and Jurafsky 2009).

Since the aim of our study is to identify and map adaptation from policy texts, we make use of a distinct class of algorithms within the general machine learning literature: artificial neural networks (ANN). ANN can be used both for supervised and unsupervised learning and have proven to be extremely effective in solving nonlinear problems with high-dimensional inputs, as for speech and image recognition, natural language processing, and text mining. ANN are a form of artificial intelligence, based on function approximations, that have been extensively used for ML.

Introductions to ANN often follow the analogy of the human brain, as in many ways ANN have a similar structure. ANN consist of several layers of artificial neurons (nodes), as depicted in Fig. 1. ANN consist of one or several layers, whereby each layer output is the input for the next layer. When more than three layers are used, one often speaks of 'deep learning', which is effective for high-dimensional nonlinear learning, such as image recognition. The computations happen in the nodes: a node receives an input and combines this with weight coefficients that amplify or dampen that input, and consequently is it passed through a non-linear activation function. ANN learn to classify inputs from training examples with known labels in a training process that minimizes the errors of ANN outputs against the known labels by adjusting node weights. In the next section, we present how we operationalize our single-layer ANN for this study.

Fig. 1 An example artificial neural network with three layers. The input layer receives a vector of $n$ inputs ($x_1, x_2, ..., x_n$), and the output layer produces a single value ($Y$). There is a single hidden layer of $p$ neurons ($P_1, P_2, ..., P_p$). Neurons are connected via synapses, shown as lines

## Creating and training the model

### Problem definition and conceptualization

In this article, we aim to explore machine learning as a method to assist researchers in identifying climate change adaptation actions across government by using policy documents as data. We assume here that reference to adaptation in policy texts is an early signal of adaptation being integrated in a specific department or organization, and as such could guide researchers to explore these processes of integration further.

To do so, we develop a ML model able to classify textual passages (hereafter 'text blocks' or simply 'blocks') as relevant to climate change adaptation. The model will receive a block as input and should be able to produce a label to classify it as 'Adaptation', 'Mitigation', or 'Non-climate' with a certain probability. To produce such a model, we first select an appropriate corpus of documents for training that has been annotated by experts. Then, we pre-process and clean the documents, transform them to extract appropriate features, select a ML model, train it, and evaluate its performance. Model evaluation is done both by comparing model predictions against a human panel at block level and comparing model performance against data that have been annotated but not used for training using cross-fold validation. Once a satisfactory performance of the model has been achieved, we interpret the patterns learned and apply them for further decision-making in a climate change adaptation context. In the remaining sections, we detail the above-mentioned process.

### Case selection

For our case study, we use policy texts from the UK as data for three reasons. First, since the UK is among the forerunners globally when it comes to climate change adaptation (Lesnikowski et al. 2016), there is a substantive body of policy texts to create a reasonably large corpus for training and testing the algorithm. Second, the UK government has adopted a comprehensive understanding of adaptation (in contrast to a more narrow, sectoral focus on adaptation in other countries), with the UK Climate Change Act (2008) forcing the ministries, departments, and statutory bodies to report on their climate change risks and adaptation actions every 5 years (Turnpenny et al. 2005; Howarth et al. 2018; Massey and Huitema 2012; Tompkins et al. 2010, Lorentz et al., 2019). This increases the likelihood that empirical evidence on policy integration is present in other sectors than the leading UK departments: the Department of Environment Food and Rural Affairs (DEFRA) and the Environment Agency (EA). Third, the access to this data is relatively easy as most recent government documents are available on one online repository which offers a comprehensive data source with easy accessibility. Moreover, policy texts are available in English which is

convenient given the international composition of the research team and the NLP tools in for English texts.

### Data collection, pre-processing, and feature extraction

Figure 2 shows the workflow for collecting data and training the model. As discussed in 'Machine learning for adaptation policy research: a brief introduction' section, supervised machine learning algorithms require labelled data from which to learn and an objective to accomplish. Selection of accurate training data is essential for the performance of the algorithm. The UK website 'gov.uk' lists documents from all 25 Ministerial departments and 385 UK agencies and public bodies. We scraped all policy documents classified as 'policy paper' with PDF as file type. This corpus of documents consists of 'general policy' documents and can be assumed that it is skewed, as most documents are not relevant for climate research.
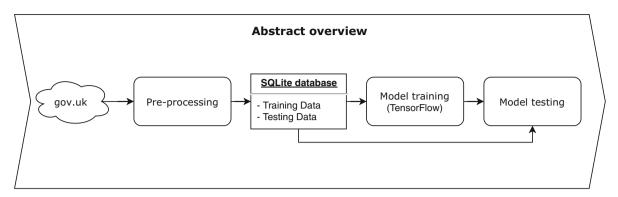
To create a training set, 10 documents were hand-picked for each of the 'Adaptation' and 'Mitigation' categories. For the 'Adaptation' category, these included the UK National Adaptation Strategy documents, UK Climate Change Risk Assessments, and reporting on key adaptation progress in the UK. For the category 'Non-climate', we hand-picked 20 documents given in the diversity of topics in this class. Documents in this category were checked for the absence of reference to climate change using the PDF search function and included the documents on, for example, pension reforms, nursing strategy, gender pay gap report, and the strategic framework to road safety. The policy papers included in our sample span a time period of 2008–2018. The longitudinal dimension allows us to track changes over time but is restricted to 10 years due to the limited availability of earlier documents in the database (the database is being continuously updated with recent and past documents). All raw data was collected in summer of 2018.

To encapsulate as much contextual information as possible to enable qualitative interpretation of the results in later stages, we aimed to analyse paragraph-level items. Within those paragraphs, we use a 'bag-of-words' approach where each paragraph is represented as a vector of word counts. For defining the training set, all paragraphs from a document were labelled with the same class as the document. While this introduces some level of feature contamination as some paragraphs could contain information on adaptation, mitigation, or even non-climate, this bottom-up approach allowed the ML algorithm to learn different interpretations and associations of what constitutes 'Adaptation' in policy texts.

Given the complex structure of the original PDF documents and the difficulty of parsing them into the original paragraph structure, we tried to approximate a paragraph structure by splitting the text on double newline characters, which
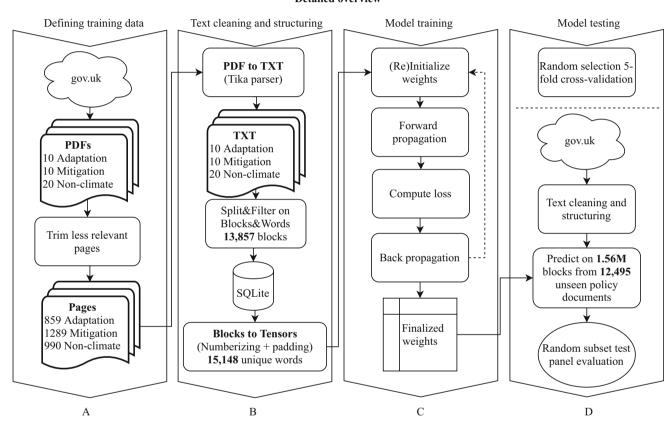
# Workflow overview



**Fig. 2** Workflow of the UK climate change adaptation case study

translates to a series of uninterrupted text separated by an empty line. We therefore refer to these as 'blocks' of text rather than paragraphs. To further clean the dataset, irrelevant pages from the PDFs were removed manually (title page, table of contents, acknowledgement, reference lists), and a minimum and maximum block length were established. The minimum block length (10 words) allowed for removing incoherent blocks of texts, for example, headers and footers and content of tables and figures. Setting a maximum block length (200 words) ensured topic specificity of the paragraph and reduced the training time of the algorithm. Data was further

pre-processed by removing word classes deemed uninformative—prepositions, determiners, conjunctions, and pronouns/prepositions using the Python module 'Natural Language Toolkit Perceptron Tagger'—and removing remaining words containing fewer than 3 or more than 20 characters. Dimensionality of the data was further reduced by removing punctuation marks and converting all uppercase letters to lowercase. Finally, a dictionary of 15,148 unique words was identified from the training set.

For fast data retrieval, a SQLite database was created to store all relevant data. Alongside the blocks, the original text

belonging to that block and additional metadata (including date of publishing and department for webscraped data) was stored for further analysis. Prediction probabilities for all three classification labels are stored in the database followed by a predicted class based on the highest probability. The final database contained 13,857 blocks of text as training data out of which 3804 were labelled 'Adaptation', 5555 'Mitigation', and 4489 'Non-climate'.

## Machine learning using an artificial neural network

For block classification, a neural network with a single hidden layer was developed using Python TensorFlow Estimator API. This API provides a simplified method for the construction of ANN, by providing high-level encapsulations for model training, evaluation, prediction, and exporting. The required input for any neural network is a uniformly shaped array of numerical values. To fulfil that requirement, unique words present in the training set were mapped to integers, and padding was added to blocks shorter than the maximum block size (set to be 200). The resulting array is called a 'tensor.' The Python TensorFlow module was used to build a 'bag-of-words' neural network model. The 'bags' in our case are counts of words in each block, labelled with their corresponding class. The hidden layer of the network is a word-embedding layer that maps words into dense vectors, called embeddings, of size $D = 50$. Word embeddings are used for learning an efficient representation in which similar words have a similar encoding. Embedding layer weights are also trainable. The output layer is comprised of three nodes, one for each class the model predicts, i.e. 'Adaptation', 'Mitigation', and 'Non-climate'. Recalculation of the loss function with the entire batch is done 100 times (number of steps) before the final weights are determined by the model.

## Evaluating the model performance

We evaluated the model performance in two ways. First, we applied 5-fold cross-validation on document level multiple times. In cross-validation, a subset (1/5th) of the entire dataset is rotationally left out for testing, and the remainder (4/5ths) is used for training. By doing cross-validation on document level, instead of block level, we minimize possible bias from the training set hidden in the writing style and vocabulary of a single author. During every fold, blocks from two documents for 'Adaptation' and 'Mitigation' and from four documents for 'Non-climate' were randomly selected and used for testing. Selected documents were disallowed from consecutive selections so that no document was tested twice. A total of 5 rotations were done to assess all documents in the training set. We repeated the process 10 times to address biases in document selection. Table 1 shows the confusion matrix averaging 10 runs of randomized selection 5-fold cross-validation. For the 'Adaptation' class, precision is 77% and recall 81%. For the 'Mitigation' class, precision is 77% and recall 81%. For 'Non-climate' precision is 82% and recall 74%. Results show an overall accuracy of 78%. Cohen's kappa statistic is 0.68 indicating substantial agreement.

To further evaluate model performance, we compared model decisions with those of a panel of five experts. A set of blocks was manually classified by experts (Graduate students from the Master Climate Studies of Wageningen University, Netherlands) and by the model. For each block, every human evaluator was able to classify it in one of the three labels or to leave a comment. The panel decision was taken by simple majority. For the ANN model, we recorded the confidence of the model predictions, which was the probability assigned to each class label. Model confidence values vary from 0.33 to 1. The ANN model decision was given to the label of the highest probability. We classified ANN model predictions as low, medium, and high confidence predictions corresponding to the intervals $(0.33 <= p < 0.6)$, $(0.6 <= p < 0.8)$, and $(p >= 0.8)$ accordingly. The comparison with the human panel was done in two phases. First, a random sample of 90 blocks was selected in such a way to include 30 blocks from each confidence interval for model predictions. In the second phase, we assessed another 85 documents where model decisions were in the high confidence range. In both phases, we assessed (a) the agreement of decisions across the human panel, (b) the agreement between the panel majority,

**Table 1** Confusion table averaging the results of 10 runs of randomized selection 5-fold cross-validation

| | | Predicted 'Adaptation' | 'Mitigation' | 'Non-climate' | Total |
|---|---|---|---|---|---|
| Actual | 'Adaptation' | 3090 (49) | 511 (39) | 203 (23) | 3804 |
| | 'Mitigation' | 580 (65) | 4423 (63) | 552 (80) | 5555 |
| | 'Non-climate' | 349 (35) | 816 (53) | 3330 (54) | 4495 |
| | *Total* | *4019* | *5750* | *4085* | *13,854* |

Standard deviation of the value over 10 runs is shown between brackets. Model predictions are shown in columns and actual labels in rows. For example, first row, second column shows that 511 blocks were on average labelled as 'Mitigation' instead of 'Adaptation', across the 10 repetitions, with a standard deviation of 39 blocks

**Table 2**    Summary of experiments evaluating ANN model performance with human panel decisions

| Phase | ANN Model Confidence Level | Number blocks | Agreement of human panel (five voters) (κ) | Agreement between ANN model and panel consensus (κ) | ANN model accuracy |
|---|---|---|---|---|---|
| Phase A | Stratified | 90 | 0.42 (Moderate) | 0.34 (Fair) | 56.6% |
| | Low (0.33–0.6) | 30 | 0.41 (Moderate) | 0.02 (Slight) | 40% |
| | Medium (0.6–0.8) | 30 | 0.39 (Fair) | 0.42 (Fair) | 60% |
| | High (0.8–1) | 30 | 0.37 (Fair) | 0.52 (Moderate) | 70% |
| Phase B | High (0.8–1) | 85 | 0.61 (Substantial) | 0.65 (Substantial) | 78.8% |

Fleiss' kappa reports the agreement among the human panel members. Fleiss' kappa and accuracy are used for evaluating agreement between ANN model and panel consensus

and the ANN model suggestions by calculating Fleiss' kappa statistic (κ) ranging from − 1 (complete disagreement) to 1 (complete agreement). Results are summarized in Table 2, and the strength of agreement is characterized according to Landis and Koch (1977).

The moderate agreement among the five human voters in phase A suggests that classifying text blocks related to climate change policy is a challenging task even for human judges. Comments of the human panellists indicated difficulties in differentiating climate adaptation policies from mitigation policies if a block was about climate change in general and differentiating climate policies from non-climate policies in cases where the subject was unclear.

The ANN was in fair agreement with the panel consensus in phase A, achieving an overall accuracy of 56% correct decisions. By further analysing phase A results, we noticed that the higher the confidence level of the ANN model, the higher is the agreement between the ANN model and the panel. Also, we observed that the ANN model and the panel was several times in disagreement about whether a block is related to climate or not, while there were few misclassifications between 'Adaptation' and 'Mitigation' classes. Policy documents contain commonly used keywords causing the model to struggle to correctly identify the topic of the policy

statement. In case of the difficulties in differentiating adaptation and mitigation, some feature contamination is to be expected as dual feature passages were not completely filtered from the training data.

In phase B, we focused on 85 random blocks that the ANN model predicted with high confidence. In this sample, the agreement among the human panellists was substantial, indicating less contextual ambiguity in these text blocks. The ANN is also in substantial agreement with the human panel, achieving an overall accuracy of 78%. Precision for the 'Adaptation' class is 72%, indicating the proportion of 'Adaptation' ANN decisions in agreement with the human panel. The recall for 'Adaptation' class is 93%, indicating the proportion of 'Adaptation' human panel decisions detected by the ANN model. The results reported per class in Table 3 suggest that correct predictions are most likely found with decisions of higher confidence and that the model does (at least partially) learn to distinguish between 'Adaptation' and 'Mitigation' classes when suggestions are made with high confidence. Recall and precision results for high confidence blocks indicate that the performance of the ANN model in detecting text blocks relevant to climate policy adaptation is very close to that of the human panel.

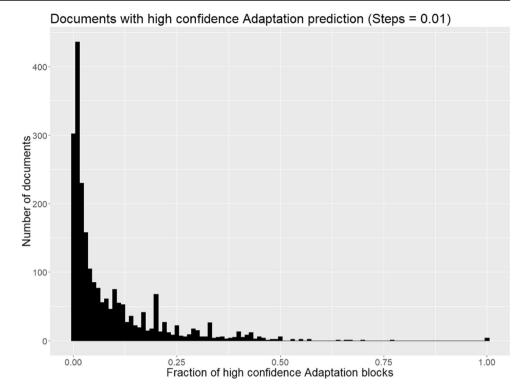**Table 3**    Class agreement between ANN model and human panel for both phases

| | | 'Adaptation' | 'Mitigation' | 'Non-climate' | Total number of blocks |
|---|---|---|---|---|---|
| Phase A | Panel predictions | 18 | 31 | 41 | 90 |
| | ANN predictions | 25 | 33 | 32 | 90 |
| | In agreement | 10 | 22 | 19 | 51 |
| | *Precision* | 40% | 67% | 59% | |
| | *Recall* | 56% | 71% | 46% | |
| Phase B | Panel predictions | 14 | 21 | 50 | 85 |
| | ANN predictions | 18 | 30 | 37 | 85 |
| | In agreement | 13 | 18 | 36 | 67 |
| | *Precision* | 72% | 60% | 97% | |
| | *Recall* | 93% | 86% | 72% | |

.The number of blocks labels assigned by both the panel and ANN are reported, along with precision and recall per class. Phase A included 90 documents across three confidence levels, while phase B includes 85 high confidence blocks

**Fig. 3** Number of documents
with high confidence (> 0.8)
adaptation predictions. The *x*-axis
represents the relative number (as
a fraction of total blocks in the
document) of high confidence
adaptation predictions within a
document. Note that we did not
remove the training and test
documents, which largely
explains the 1.0 fraction
documents



Documents with high confidence Adaptation prediction (Steps = 0.01)

## Using the algorithm to explore new UK policy documents

To illustrate the value of our trained model for climate policy research, we applied it to the larger sample of policy documents and explore the data qualitatively and quantitatively. We used the ANN model to identify high confidence (> 0.8–1.0) climate change adaptation blocks out if the corpus of 12,367 documents were classified as 'policy papers' (2008–2018) in the UK database. Processing these documents using the steps presented above (Fig. 2) resulted in 1.6 million blocks of text that was assessed by our model. This section discusses some of the illustrations of using our model for a variety of questions.

## Identifying relevant policy documents for researching policy integration

Our database contains text blocks from 12,367 policy papers related to a variety of topics, and in a manual-coding approach, most of them would not have been included. To identify the relevance to climate policy adaptation among all documents, we calculated the fraction of blocks per document that the ANN model marked as "Adaptation" with high confidence; see Fig. 3. Figure 3 verifies that adaptation is not the main topic in most documents in our database, as the majority of documents have a very low proportion of "Adaptation" blocks (i.e. fraction below 0.1). Arguably climate change adaptation policy research focuses mostly on the documents that

contain a larger fraction of blocks as this increases the likelihood that useful information is obtained. In the UK context, 1043 documents fall within this category and therefore could report useful information. When comparing over time periods, changes in the distribution of the fraction of blocks could be used as an indicator for how adaptation has become integrated across multiple departments and organizations.
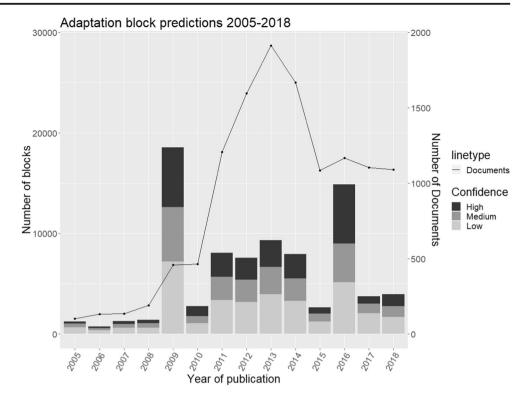
## Emergence of adaptation as a policy issue in the UK: quantitative assessment

We can use the model output as an indicator for adaptation policy issue attention and look for possible explanations for increased and decreased issue attention. The model output shown in Fig. 4 demonstrates the model predicts ~ 18,000 adaptation blocks in the year 2009, followed by a period of ~ 3000–10,000 blocks (2010–2015), a peak in 2016 (~ 15,000 blocks), and finally 2 years of ~ 4000 blocks.

The 2009 peak could be explained by several ongoing activities at domestic and international level. For example, the UK Climate Change Act was adopted in 2008, which was considered a landmark commitment to act on climate change across departments and across levels. The Act resulted in various immediate institutional innovations, including installing a Climate Change Committee with an Adaptation Sub Committee, new reporting procedures, and the 5-year UK risk assessment. In October 2008, the Prime Minister Brown also installed the new Department for Energy and Climate Change (DECC) (which was later dismantled in 2016). Since policies

**Fig. 4** Number of 'Adaptation' block predictions on the total set of policy documents obtained from gov.uk. A number of blocks are on the left *y*-axis (bars) and the number of documents published on the right *y*-axis (line)



tend to take time to get developed and adopted, it can be expected to see the effect the subsequent year. Internationally, the COP15 ('Copenhagen Summit') in 2009 raised domestic political attention to climate change as it was expected that a new climate agreement would be signed with a strong focus on adaptation. In the same year, the EU adopted a green paper entitled "Adapting to climate change: towards a European framework for action" in which the commission proposed a number of EU wide activities to create an EU wide strategy on climate change adaptation and to ensure timely actions of EU member states, including the UK (Biesbroek and Swart 2019). These activities could explain the rapid increase of predicted high confidence adaptation blocks.

The lower frequency of blocks for the period between 2009 and 2015 can be the result of implementing some of the proposed actions, although some literature suggests that the Act and associated actions in 2008–2009 did not necessarily translate to concrete measures being implemented on the ground due to a variety of socio-economic and political reasons (for more details see, e.g. Lockwood (2013)). The 2016 peak could possibly be explained by the second call for evidence and reporting requirement under the UK 2008 climate change act.

### Shifts in departmental reporting on adaptation over time

By collecting meta-data of the governmental organizations that submitted policy documents to the UK repository, we can look more closely at horizontal policy integration, i.e. between different departments and organizations. Two governmental organizations, DEFRA ($n = 10,831$) and the Environment Agency ($n = 9689$), have reported most blocks over time. We observe a shift from most blocks originating from DEFRA (2008–2013) to the Environment Agency (2014–2018). When looking at the percentage of high-frequency blocks for each department, we see a peak of reporting in 2016, which could possibly be explained by the required reporting under the UK climate change act. Overall, the departments with the highest average number of all confidence level adaptation blocks after DEFRA and the EA for the period 2005–2018 are the department of Department for Transport (1032 blocks), High Speed Two (HS2) Limited (548 blocks), and Cabinet Office (530 blocks)(Fig. 5).

### Qualitative assessment of policy integration

The resulting database also allows for qualitative investigations of policy integration. To illustrate this, we randomly sampled 200 high confidence adaptation blocks from those documents with a low fraction of these blocks (fraction < 0.01%, see Fig. 3). We report on three examples to demonstrate the range of possible qualitative assessments from the data collected.

The first example is where the model predicts with high confidence (0.945) block as adaptation. The extracted block is from a report from the UK Home Office (2011) entitled: 'The United Kingdom's Strategy for Counter Terrorism'. The report presents how the UK is planning for the possible acts of
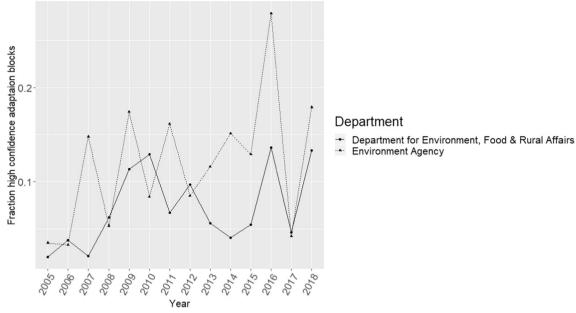
**Fig. 5** High confidence adaptation predictions for the two departments with the highest climate adaptation output as fraction of their total number of blocks in the dataset: DEFRA and EA.

domestic and international terrorism, as well as possible strategies to navigate these risks. Our model extracted one paragraph which points to synergies between planning and taking actions to adapt to natural hazards and climate change and reducing the risks from and responding to terrorist attacks:

> These plans focus on flooding and other natural hazards, in response to the independent review by Sir Michael Pitt after floods in 2007. But some capabilities (alerting systems, mechanisms for cooperation with emergency responders, contingency and business continuity) will also help prepare critical infrastructure to respond to terrorist attacks.

The excerpt above shows that the UK Home Office, which is generally not included in UK studies on climate change adaptation, is considering the linkages between natural hazards, climate change, and terrorism as early as 2011.

The second example is from a report of the Department of Transport (2008) entitled *Roads: Delivering Choice and Reliability* where the model predicted with high confidence (0.988) that the following paragraph is climate relevant:

> We are already doing much to reduce the impact of these conditions. For example the Highways Agency has already improved drainage and road surface standards to increase resilience. We will continue to take steps to ensure that our infrastructure is planned, designed, maintained and managed to be resilient to future climate

impacts, through the application of tools such as the Highways Agency's climate change adaptation strategy.

In this excerpt, reference is made to concrete policy tools and actions that are important for analysts to understand when investigating how the UK transport sector was starting to plan for adapting their infrastructure to climate change in 2008. Such information could be useful to track progress in achieving these climate policy goals and evaluate the adequacy and effectiveness of policy tools used and actions taken by comparing subsequent policy texts.

Our third and final example is to showcase that this approach also allows us to identify relevant subnational information, including reference to specific investments in regional climate change adaptation. In this particular example, the model identified the following block from the "Second Report of Session 2016–17, Flooding: Cooperation across Government" report (2016) to be of relevance (0.915):

> Many communities in the North were badly affected by flooding this winter. As part of the government's £700 million boost to flood defence and resilience spending, £150 million will be invested in flood defence schemes in Leeds, Cumbria, Calder Valley and York, which will better protect 7,400 properties. The government will also invest up to £25 million in flood defences in Carlisle once the Environment Agency has concluded a review of its needs, and will provide funding to support delivery of the final phase of the Leeds Flood Alleviation Scheme in later years subject to business case approval.

# Discussion: the value of machine learning for adaptation policy integration research

In this article we explored the use of machine learning methods for the study of climate change adaptation policy. By applying a neural network model to the UK case, we were able to identify a number of potentially relevant blocks of text extracted from existing policy documents. The model allowed us to explore the patterns of policy attention across departments and over time and identified relevant sections of text from policy texts that are usually not included in policy integration studies. We argue that such tools can assist research on adaptation policy to quickly process large volumes of policy text and take a data-driven approach to study policy integration comprehensively.

When comparing our empirical results to other studies conducted on UK climate change adaptation policy landscape, we find complementary results. Lorenz et al. (2019), for example, use an extensive content analysis of 146 policy documents (2006–2015) to identify 568 different actors involved in the adaptation policy landscape in the UK. Although our intention was not to identify actors, our inventory of key national governmental actors involved is quite similar. Like the previous studies in the UK (Turnpenny et al., 2005, Tompkins et al. 2010), they recognize the challenges of what gets counted as adaptation in the coding, which data sources to use to extract relevant information, and what information to read. We argue our method could be a useful first step in data collection and processing as it largely tackles these issues, potentially allowing for a more comprehensive assessment of the adaptation policy landscape.

Machine learning methods, including the model we developed in this article, arguably create new and more elaborate empirical and theoretical questions that can only be answered comprehensively with such methods, for example, about how the adaptation policy landscape evolves and what drives these changes. Such methods have the potential to significantly alter the way adaptation policy research is being conducted, with a stronger focus on combining qualitative and quantitative methods to answer these crucially relevant societal questions. But new methods are also vulnerable to criticism, for example, about the validity of the method and its findings, the usefulness of its application, and the necessary skills needed to implement it. The purpose of this article is not to silence these criticisms but rather to demonstrate the potential of these tools and, in doing so, to enrich the set of tools being used in climate policy research and policy advice.

We are the first to admit that our model is far from perfect, and several steps could be taken to further improve it to have higher prediction accuracy. We consider these to be important lessons for further developing these methods. Most importantly, the problem formulation is important in defining metrics for ML model evaluation and training. In this work, we considered that each block was essentially assigned to a single label. However, the ANN mode we used was able to handle the problem as a multi-label classification, i.e. produce probabilities next to each label. Future research can investigate further this direction. Second, we purposefully decided not to hand-code passages for training purposes as this would undermine our initial idea of having a model learning from policy documents what adaptation is rather than predetermining what adaptation is (and is not). It would be worthwhile to explore the added value of making use of hand-coded blocks to explore the impact it has on model accuracy. Blocks coded by the student panel accompany this paper as open data and could be used for further investigating this research direction. Third, while we have taken a number of key steps to pre-process the scraped files, working with PDFs means there will be problems with the quality of the data. Although alternatives exist in the UK context (HTML versions of the text), we decided to accept some inaccuracies in our model as PDFs are likely the most frequently used type of document for policy research. Fourth, we had limited data for training as adaptation is still a relatively new policy field, even in forerunner countries such as the UK. To further improve machine learning for creating artificial intelligence models to capture the various dimensions of adaptation in policy documents, a more elaborate model with a larger quantity of training data is required. It will only be a matter of time before more training data becomes available to further advance our model. For the purpose of this article, we focused on a subset of policy documents (classified as 'policy papers'), but there is a large volume of documents we did not assess for practical reasons including processing time and data storage. Upscaling this method to include more data sources (i.e. include different types of documents) and/or country contexts would require incorporation of big data methods to enable scalable processing.

Upscaling of our approach to other country contexts remains challenging. First and foremost, it only works with digitized data, which in many parts of the world is a severe limitation. Records of scanned images of policy texts can be used through text recognition software, but such processes typically introduce new types of errors. Poor digitalization is always a limitation, but given increased digitization of societies across the globe, the usefulness of these algorithms should rapidly increase. Second, not all data is as neatly accessible as is the case of the UK. In many cases, alternative methods should be used such as scraping multiple websites to create databases. Finally, although there is an advantage to creating contextually sensitive understandings of what adaptation through tailored algorithms, it requires training and testing of the algorithm in each context for which it is being used. This is time and resource intensive, but by no means as intensive as manual assessments.

In terms of next steps, we can further explore the database created, for example, using topic modelling to extract latent structures from the blocks and get a sense of the key themes

and topics discussed and how they might have changed over time and vary between departments and public sector organizations. See Lesnikowski et al. (2019) for an example of using topic modelling for climate change adaptation research.

Although our efforts here are exploratory in nature and limitations exist, the findings of this article point to the added value of using machine learning as a new way of informing the monitoring and evaluation of countries and global assessments such as the global stocktake on climate change adaptation (Berrang-Ford et al. 2019). These—or similar—models can be used as an indication of whether or not countries are adapting and can function as a parallel process to assess whether countries are moving beyond symbolic reporting under the UNFCCC reporting requirements through their National Communications and Nationally Determined Contributions by assessing adaptation in all the policies becoming digitally available.

## Compliance with ethical standards

**Conflict of interest**    The authors declare that they have no conflict of interest.

## References

Anastasopoulos, L.J., Whitford, A.B., (2019). Machine learning for public administration research, with application to organizational reputation. *J Public Adm Res Theory*

Bauer A, Feichtinger J, Steurer R (2012) The governance of climate change adaptation in 10 OECD countries: challenges and approaches. J Environ Policy Plan 14:279–304. https://doi.org/10.1080/1523908X.2012.707406

Bauer A, Steurer R (2014) National adaptation strategies, what else? Comparing adaptation mainstreaming in German and Dutch water management. Reg Environ Chang 15:341–352. https://doi.org/10.1007/s10113-014-0655-3

Biesbroek R and Candel JLL (2020) Explanatory mechanisms for policy (dis)integration: food policy and climate change adaptation policy in the Netherlands, Policy Sciences, 53, p61–84. https://doi.org/10.1007/s11077-019-09354-2

Biesbroek R, Swart RJ 2019. *Adaptation policy at supranational level? Evidence from the European Union*, in (ed), in: research handbook on climate change adaptation policy. Edward Elgar Publishing, UK

Biesbroek R, Berrang-Ford L, Ford JD, Tanabe A, Austin SE, Lesnikowski A (2018) Data, concepts and methods for large-n comparative climate change adaptation policy research: a systematic literature review. Wiley Interdiscip Rev Clim Chang 9:e548. https://doi.org/10.1002/wcc.548

Berrang-Ford L, Biesbroek R, Ford JD, Lesnikowski A, Tanabe A, Wang FM, Chen C, Hsu A, Hellmann J, Pringle P, Grecequet M, Amado J-C, Huq S, Lwasa S, Heymann SJ (2019) Tracking global climate change adaptation among governments. Nat Clim Chang 9:440–449. https://doi.org/10.1038/s41558-019-0490-0

Candel JJL, Biesbroek R (2018) Policy integration in the EU governance of global food security. Food Sec 10:195–209. https://doi.org/10.1007/s12571-017-0752-5

Candel JLL, Biesbroek GR (2016) Toward a processual understanding of policy integration. Policy Sci 49(3):211–231. https://doi.org/10.1007/s11077-016-9248-y

Cejudo GM, Michel CL (2017) Addressing fragmented government action: coordination, coherence, and integration. Policy Sci 50:745–767. https://doi.org/10.1007/s11077-017-9281-5

Clar C, Steurer R (2019) Why popular support tools on climate change adaptation have difficulties in reaching local policy-makers: qualitative insights from the UK and Germany. Environ Policy Gov 28:172–182. https://doi.org/10.1002/eet.1802

Cody EM, Reagan AJ, Mitchell L, Dodds PS, Danforth CM (2015) Climate change sentiment on twitter: an unsolicited public opinion poll. PLoS One 10:e0136092. https://doi.org/10.1371/journal.pone.0136092

Creutzig F, Lohrey S, Bai X, Baklanov A, Dawson R, Dhakal S, Lamb WF, McPhearson T, Minx J, Munoz E, Walsh B (2019). Upscaling urban data science for global climate solutions. Global Sustainability 2. https://doi.org/10.1017/sus.2018.16

Donner SD, Kandlikar M, Webber S (2016) Measuring and tracking the flow of climate change adaptation aid to the developing world. Environ Res Lett 11:054006. https://doi.org/10.1088/1748-9326/11/5/054006

Dupuis J, Knoepfel P (2013) The adaptation policy paradox: the implementation deficit of policies framed as climate change adaptation. Ecol Soc 18(4):31. http://dx.doi.org/10.5751/ES-05965-180431

EEA (2014) National adaptation policy processes in European countries — 2014. European Environment Agency, Copenhagen

Farrell J (2016) Network structure and influence of the climate change counter-movement. Nat Clim Chang 6:370–374. https://doi.org/10.1038/nclimate2875

Ford JD, Tilleard SE, Berrang-Ford L, Araos M, Biesbroek R, Lesnikowski AC, MacDonald GK, Hsu A, Chen C, Bizikova L (2016) Opinion: big data has big potential for applications to climate change adaptation. PNAS 113:10729–10732. https://doi.org/10.1073/pnas.1614023113

Grimmer J (2015) We are all social scientists now: how big data, machine learning, and causal inference work together. PS: Political Science & Politics 48:80–83. https://doi.org/10.1017/S1049096514001784

Henstra D (2017) Climate adaptation in Canada: governing a complex policy regime. Rev Policy Res 34:378–399. https://doi.org/10.1111/ropr.12236

Howarth C, Morse-Jones S, Brooks K, Kythreotis AP (2018) Co-producing UK climate change adaptation policy: an analysis of the 2012 and 2017 UK climate change risk assessments. Environ Sci Pol 89:412–420. https://doi.org/10.1016/j.envsci.2018.09.010

Hobson L Hannes H, Cole H (2019) Natural language processing in action. Manning. ISBN: 1617294632

IPCC (2014). Summary for policymakers., in: Field, C.B. (Ed.), Climate change 2014: impacts, adaptation, and vulnerability. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 1–32

Krippendorff K (2018). Content analysis: an introduction to its methodology, 4th ed. Sage publications

Landis J, Koch G (1977) The measurement of observer agreement for categorical data. Biometrics 33(1):159–174. https://doi.org/10.2307/2529310

Lazer D, Pentland A, Adamic L, Aral S, Barabási A-L, Brewer D, Christakis N, Contractor N, Fowler J, Gutmann M, Jebara T, King G, Macy M, Roy D, Van Alstyne M (2009) Computational Social Science. Science 323(5915):721–723. https://doi.org/10.1126/science.1167742

Lenschow A, Bocquillon P, Carafa L (2018) Understanding coherence between policy spheres. Environ Policy Gov 28:323–328. https://doi.org/10.1002/eet.1818

Lesnikowski A, Belfer E, Rodman E, Smith J, Biesbroek R, Wilkerson JD, Ford JD, Berrang-Ford L (2019) Frontiers in data analytics for adaptation research: topic modeling. Wiley Interdiscip Rev Clim Chang e576. https://doi.org/10.1002/wcc.576

Lesnikowski AC, Ford J, Biesbroek R, Berrang-Ford L, Heymann SJ (2016) National-level progress on adaptation. Nat Clim Chang 6: 261–264. https://doi.org/10.1038/nclimate2863

Lesnikowski AC, Ford JD, Berrang-Ford L, Barrera M, Heymann J (2015) How are we adapting to climate change? A global assessment. Mitig Adapt Strateg Glob Chang 20:277–293. https://doi.org/10.1007/s11027-013-9491-x

Lockwood M (2013) The political sustainability of climate policy: the case of the UK climate change act. Glob Environ Change 23:1339–1348. https://doi.org/10.1016/j.gloenvcha.2013.07.001

Lorenz S, Porter, JJ, Dessai, S (2019) Identifying and tracking key climate adaptation actors in the UK. Reg Environ Change 19(7):2125–2138. https://doi.org/10.1007/s10113-019-01551-2

Lynam T (2016). Exploring social representations of adapting to climate change using topic modeling and Bayesian networks. Ecol Soc 21

Maor M, Tosun J, Jordan A (2017) Proportionate and disproportionate policy responses to climate change: core concepts and empirical applications. J Environ Policy Plan 19(6):599–611. https://doi.org/10.1080/1523908X.2017.1281730

Martin JH, Jurafsky D (2009). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition*: Pearson/prentice hall upper Saddle River

Massey E, Huitema D (2012) The emergence of climate change adaptation as a policy field: the case of England. Reg Environ Chang 13: 341–352. https://doi.org/10.1007/s10113-012-0341-2

Peters BG (2015) Pursuing horizontal management. The politics of public sector coordination. University Press of Kansas, USA

Runhaar H, Wilk B, Persson Å, Uittenbroek C, Wamsler C (2018) Mainstreaming climate adaptation: taking stock about "what works" from empirical research worldwide. Reg Environ Chang 18:1201–1210. https://doi.org/10.1007/s10113-017-1259-5

Russell, S. J., & Norvig, P. (2009). Artificial intelligence: a modern approach, Pearson Education Limited

Tompkins EL, Adger WN, Boyd E, Nicholson-Cole S, Weatherhead K, Arnell N (2010) Observed adaptation to climate change: UK evidence of transition to a well-adapting society. Glob Environ Chang 20:627–635. https://doi.org/10.1016/j.gloenvcha.2010.05.001

Tosun J, Lang A (2017) Policy integration: mapping the different concepts. Policy Stud 38:553–570. https://doi.org/10.1080/01442872.2017.1339239

Turnpenny J, Haxeltine A, Lorenzoni I, O'Riordian T, Jones M (2005) Mapping actors involved in climate change policy networks in the UK. Tyndall Centre Working Paper No. 66. Tyndall Centre for Climate Change Research: 1–18

Uittenbroek, C.J., Mees, H.L.P., Hegger, D.L.T., Driessen, P.P.J., 2019. The design of public participation: who participates, when and how? Insights in climate adaptation planning from the Netherlands. J Environ Plan Manag 0, 1–19

Wellstead A, Stedman R (2014) Addressing the challenges of adaptation to climate change policy: integrating public administration and public policy studies. Int J Public Adm 37:999–1010. https://doi.org/10.1080/01900692.2014.907313

Zizka J, Darena F, Svobodja A (2020) Text mining with machine learning. Principles and techniques. CRC press. Tailor and Francis group