This work was written as part of one of the author's official duties as an Employee of the United States Government and is therefore a work of the United States Government. In accordance with 17 U.S.C. 105, no copyright protection is available for such works under U.S. Law.

Public Domain Mark 1.0
https://creativecommons.org/publicdomain/mark/1.0/

Access to this work was provided by the University of Maryland, Baltimore County (UMBC) ScholarWorks@UMBC digital repository on the Maryland Shared Open Access (MD-SOAR) platform.

Please provide feedback

Please support the ScholarWorks@UMBC repository by emailing scholarworks-group@umbc.edu and telling us what having access to this work means to you and why it's important to you. Thank you.

Journal of Air Transport Management 102 (2022) 102228

Contents lists available at ScienceDirect

Journal of Air Transport Management

journal homepage: www.elsevier.com/locate/jairtraman

Using sentiment analysis to reinforce learning: The case of airport community engagement[1]

Tony Diana

Federal Aviation Administration, 800 Independence Avenue, SW Washington, DC, 20591, United States

1. Introduction

Community engagement has become critical for airport management and air traffic control regulators. The success of airport development programs often depends on the support of airport community residents. To maintain collaboration among all these stakeholders, airport and regulator officials need to understand residents' sentiment to anticipate issues and concerns before they turn into problems that can stall important capital projects. It is not unusual to see runway construction projects run over decades and eventually get canceled.

This study uses the example of a large hub airport in the Northeast of the United States. The case is representative of the issues and problems that airport community residents experience in their engagement with the airport authority, the regulators, elected officials, and community leaders. The article illustrates how residents' sentiment expressed in digital prints can be leveraged to inform decision-making in community engagement. As agents face an uncertain environment, they must re-assess their strategy with little knowledge of the magnitude of benefits or losses related to their policy implementation. The literature on Reinforcement Learning often emphasizes the agent's choice between 'exploiting' a situation or 'exploring' new alternatives (Sutton and Barto, 2018; Winder, 2021).

This article establishes a link between two of the fastest-growing areas of Machine Learning (ML), which is a division of Artificial Intelligence (AI): Natural Language Processing (NLP) and Reinforcement Learning (RL). Humans use natural language in the forms of text, speech, and sign as opposed to programming language designed for computers. Sentiment analysis as part of NLP includes the study of sentiment, opinion, intent, and emotion to explain attitude and behavior. RL focuses on how agents (humans or entities) learn from their environment to derive rewards. The digitization of newsprints and social media has made it possible to take advantage of sophisticated ML algorithms, which can provide insight into how agents behave and select courses of action.

After defining the concepts of sentiment analysis and community engagement, we will delve into the 'multi-armed bandit' (MAB) problem and the concept of 'contextual multi-armed bandit' (CMAB) that adds the significance of the environment in an agent's strategy selection. Then, we will apply a classification model to determine how the selected features described in a later section can predict the choice of one of three policy categories:

- Exploitation: taking advantage of implemented policies and do not change them,
- Exploration: innovating and experimenting with new policies to broaden community stakeholders' support and attempt to reduce dissatisfied residents,
- Pause: putting on hold any further action to evaluate the impact of implemented policies and determine future directions.

The topic of community engagement and outreach is of great interest to aviation practitioners because agents as individuals or organizations must often make important decisions in a context of uncertainty. Sentiment can shift unexpectedly due to circumstances that may be beyond agents' control. This article offers a unique perspective on the relationship because no research to the author's knowledge has explored the interactions between sentiment analysis and RL in the context of airport community engagement. Diana (2021) studied how sentiment analysis could help anticipate residents' attitudes based on changes in key operational factors. It showed that operational factors such as the use of specific runway configuration, ceiling, and visibility could predict variations in the sentiment of airport communities.

Secondly, this article proposes a methodology to select a policy based on a classification model. The approach can be replicated by other airports or regulators in managing community involvement.

2. The context of community engagement

This paper examines how sentiment analysis may reinforce agents' learning as they interact with airport communities' stakeholders.

[1] This article is the result of independent research. No funding was provided.
E-mail address: tony.diana@faa.gov.

https://doi.org/10.1016/j.jairtraman.2022.102228
Received 22 June 2021; Received in revised form 21 February 2022; Accepted 9 May 2022
Available online 24 May 2022
0969-6997/Published by Elsevier Ltd.

Engagement involves policies and actions that agents implement to improve communication and messaging to airport community residents, such as workshops, formal briefings, press releases, conferences, and mailing.

Together with 'environment' and 'policy,' 'agent' represents a key component in the RL framework. The main purpose of RL is to establish how agents can maximize their rewards/benefits (or minimize their regrets/losses) by learning from interactions with their environment.[1] Agents such as airport management or regulators learn through a series of successes and failures in their interactions with community residents. Based on their experience, agents select among competing courses of action to maximize their expected 'rewards' or minimize their 'regrets.'

However, agents may find it difficult to predict whether they can 'exploit' past policy successes, 'explore' new policy directions, or 'pause' to re-evaluate their engagement policy toward community residents. Regularly tracking the sentiment of community residents through digital prints and social media can minimize. A regular assessment of residents' attitudes and emotions can help agents be more initiative-taking in responding to unexpected changes. This explains why sentiment analysis has become an effective tool for strategy design, policy monitoring, program evaluation, and decision making. At this point of the discussion, it is important to explain how sentiment analysis can reinforce agents' learning.

First, sentiment represents an intermediate state between feeling (undefined attitude on specific topics) and opinion (more definite perspective on specific issues and problems). When sentiment becomes opinion, agents must expend more energy and resources in communication and messaging to change people's minds.

Second, sentiment analysis can effectively help agents quickly adjust policies, initiatives, or actions to environmental change. It serves to classify sentiment usually into three categories such as 'negative,' 'neutral,' or 'positive.' Sentiment analysis can strengthen both policy monitoring and evaluation. Whereas policy monitoring is rooted in facts, policy evaluation is embedded into both facts and judgment (Dunn, 2015).

Third, sentiment analysis allows agents to grasp the manifest (expressed openly in media) and latent (identified through text and semantic analysis) concerns among airport community stakeholders. Topic modeling techniques such as Latent Dirichlet Allocation, Latent Semantic Analysis, and Probabilistic Latent Semantic Analysis are algorithms that serve to cluster words in collected documents into topics that summarize key areas of preoccupation or concern for airport community residents.

Fourth, sentiment analysis can determine whether residents may 'exit,' raise 'voice,' or remain 'loyal.' Residents can decide to follow one of the three strategies that Hirschman (1970) described in his seminal book *Exit, Voice, and Loyalty*. They can 'exit,' that is, refuse to stay engaged in a dialogue with agents. This will eventually deprive agents of valuable feedback to address dissatisfaction among stakeholders. Residents can 'voice' their concerns, which alerts agents of potential obstacles to their policy and the need for change. Finally, residents can remain 'loyal' or supportive of agents' current policies.

Fifth, agents often find it difficult to measure the benefits and costs of policies and evaluate community residents' satisfaction with their policies. Moreover, not all stakeholders may agree with policy evaluation criteria.[2] Communities around airports are more likely to deal with externalities, that is, benefits or costs whose residents are enjoying or incurring without being parties to a transaction (see Coase, 1960; Baumol and Oates, 1988). Externalities can be either positive (i.e., jobs growth generated by airport activities) or negative (i.e., noise and carbon emissions generated by aircraft operations). Without considering residents' sentiment, agents are challenged to predict future courses of action without any tangible baseline. As a result, they may get surprised that community residents may be demanding to revert past program implementations and resort to legal actions.

Finally, methods such as contingent valuation to estimate use and non-use values may not always be effective to reveal preferences. Contingent valuation (CV) methods resort to 'stated preferences' methods (Bateman et al., 2002). However, CV methods can be biased because survey respondents do not have an incentive to reveal their true preferences. CV is likely to provide a biased perspective of what respondents say they will do or prefer as opposed to what they are observed to do. In some situations, the conclusions from surveys and interviews are no longer valid by the time they are published as sentiment evolves fast. Similarly, panels are subject to participants' fatigue and, as a result, their utility can steadily decline. Sentiment analysis is more effective at revealing preferences through publicly available data collected from digital print reports, social media, blogs, and vlogs.

3. Methodology

This study presents the case of a U.S. East Coast airport from 2015 to 2021. The airspace redesign program at the case study airport started in 2016. As a result, this study provides a pre- and post-implementation perspective on how the implementation influenced public sentiment and support. The next section details the analytical steps.

1. **Identification of Research Keywords.** The collection of a sample depended on the selection of keywords that described the scope and purpose of the analysis. The keywords included 'airport and aircraft noise,' 'air pollution,' 'airport community meetings and workshops,' 'flight paths and procedures,' 'airspace redesign,' and 'airport community residents' complaints.'
2. **Sample Creation.** Text data were collected from digital prints, social media, video transcripts, and press releases published by local, state, and federal politicians who participated in airport community engagement. The sample incorporated the following content: (1) publicly available digital prints and blogs scraped from websites, (2) the transcripts of videos from local news reports on issues related to airport community involvement, and (3) social media posts on Facebook and tweets from Twitter. Digital prints refer to articles published in state and local newspapers that reported on airport community events. We added social media to the digital prints data for a more comprehensive assessment of the airport community's sentiment. Digital prints articles usually offer a better balance between facts and opinions, whereas social media provide the spontaneous emotional reactions of airport community residents to events.

The document search criteria pertained to (1) aircraft noise issues related to changes in arrival and departure procedures, (2) airport/resident community involvement in workshops and meetings, and (3) federal, state, and local politicians' involvement in aircraft/aircraft

[1] Readers interested in an in-depth literature review on the various types of RL models are referred to Salvador et al. (2020).

[2] Policy makers may use performance criteria such as efficiency, effectiveness, adequacy, equity, responsiveness, and appropriateness. Others may use benefit/cost principles such as (1) the Kaldor-Hicks criterion (those who benefit in gains in efficiency compensate the losers); (2) the Rawls criterion (when a social state provides a gain in welfare for those who are worse off); (3) the Pareto criteria (when a policy creates a state in which at least one person is better off and no one worse off).

noise complaints. Digital prints were manually scraped from the Web or via software such as BeautifulSoup.[3] The content of each extracted piece was validated to make sure it pertained to the search criteria.

3. **Data Processing.** Texts are unstructured data. They need to be cleaned for punctuation, stopwords, upper cases, and digits before being tokenized. Tokenized words are reduced to their root through the process of lemmatization to be vectorized.
4. **Input Features.** For each sampled year, the model described later in this section included the following features:
   - The percentage of angry, fearful, sad, and surprise emotions detected in each sample,
   - Cumulative regrets derived from 1000 simulations over 30 periods,
   - Cumulative rewards computed as cumulated regrets, and
   - The percentage of positive, neutral, and negative sentiment.
5. **Sentiment Analysis.** The cleaned data served to derive the percent of positive, neutral, and negative sentiment, the categories of intent, and emotion detected in the sample. Sentiment and emotions represented features in the classification models. Sentiment, emotions, and intent were derived from the Komprehend[4] Artificial Intelligence application programming interface (API), as well as the Python libraries including TextBlob,[5] VADER (Valence Aware Dictionary and sEntiment Reasoner),[6] and text2emotion.[7]

Sentiment analysis refers to the extraction of text, video, or sign data (emoticon) to measure how people feel. Sentiment analysis can leverage either lexicon-based or machine-learning-based algorithms. In the former type, a document is classified based on the counts of words that match the positive or negative words in a lexicon. With the former type, we could identify subjectivity and intent using TextBlob and text2emotion. In the latter type, a classification model predicts whether a label belongs to a specific class as in the case of Komprehend API by leveraging neural networks.

6. **The Context of Airport Community Involvement.** Contextual Multi-Armed Bandit is a special case of the Multi-Armed Bandit problem. MAB refers to the slot machine set in a casino where a 'bandit' has the possibility of pulling down the lever of machines in search of maximizing rewards. Bandits do not know the distribution of rewards when they pull the lever of a slot machine. They face the problem of identifying the best set of actions without knowing in advance which one would minimize cumulative regrets.

The MBA problem can be applied to any situation where learning is sequential, and the environment is uncertain. Robbins (1952) described the MAB problems as decision-making situations when the environment is uncertain and the outcomes of actions unknown. The MAB problem has been used in advertising (Chapelle et al., 2014), website optimization (White, 2012), clinical trials (Villar et al., 2015), recommending systems, and information retrieval applied to healthcare and finance (Bouneffouf and Rish, 2019).

In the context of a community engagement, agents cannot always anticipate the reaction of residents after an engagement event. As Palmas et al. (2020:424) stated, "The problem […] boils down to the design of a learning strategy where the player needs to explore what possible reward values each slot machine can return and from there, quickly identify the one that is most likely to return the greatest expected reward." Sentiment analysis can provide the information that agents need to learn and react quickly.

The MAB problem is characterized by four key elements:

- K policies to choose from (the arms of the slot machines). Our reinforcement learning framework identified 'exploit,' 'explore,' and 'pause' as policy alternatives.
- A central agent or group of agents who pull the arms (i.e., airport authority, local, state, and federal elected officials, community advocates, and air traffic control regulators).
- Each arm has a specific probability distribution, which is unknown to agents.
- The goal is to select the arm that minimizes regrets or maximizes rewards over intervals called 'horizon.' In our case, we have three levers ('exploit,' 'explore,' and 'pause'). We ran one hundred simulations in thirty periods as the horizon.

The reward associated with each agent or 'bandit' who pulls the arm follows a Bernouilli distribution characterized as follows:

$$P_x = \binom{n}{x} p^x q^{n-x} \tag{1}$$

where x = the number of times, $\binom{n}{x}$ = the number of combinations, p = the probability of success in a trial, q = the probability of failure on a single trial, and n = the number of trials.

MAB problems mention the concepts of 'reward' and 'regret.' Bubeck and Cesa-Bianchi (2012) provided a detailed exposition of regrets in the MAB problems. If $X_{i,n}$ is a random variable for $1 \leq i \leq i$ and $n \geq 1$ (n being the number of plays and i the index that identifies the arm and $T_i(n)$ is the number of times the lever i is played in the first n plays), then regret is defined as

$$R_T = T\mu^* - \mu_{j(T)} \sum_{k=1}^{K} E(T_k(T)) \tag{2}$$

with μ^* = maxi = 1, …, k μ_j and $E(T_k(T))$ is the expectation about the number of times the policy will play machine k. Regret is "the difference between the maximum possible gain (having thrown the best lever n times, by definition the one that returns μ* as a reward) and the actual gain" (Ciaburro, 2019: 77).

To understand how agents select a policy, it is important to consider two functions. First, the *state-value function* is a matrix that represents the reward associated with a given state. The state-value function is defined as

$$V^{\pi}(s) = \sum_{t=0}^{\infty} \gamma_i r_{t+1} \tag{3}$$

where π = a policy associated with s = the state and a = the action $\pi(s,a)$, r = the reward, and γ = a discount factor. Action is the possible move of an agent, whereas states are the observations from the environment.

Second, the *action-value function* is a matrix that provides the reward for every state-action pair. The action-value function is defined as

$$Q(s,a) = r(s,a) + \gamma V^{\pi}\, \delta(s,a) \tag{4}$$

where $\delta(s,a)$ is the function that determines the new state generated by the pair (s,a). "The state value function contains the value of reaching a certain state, while the action-value function contains the value for choosing an action in a state" Ciaburro (2019: 81).

In the contextual multi-armed bandit framework, the state becomes "a description of the environment that the agent can use to carry out targeted actions" (Ciaburro, 2019:86). The context depends on the sentiment of the community residents (measured as 'negative,' 'neutral,' or 'positive' attitude in this study). In airport community engagement,

[3] See https://pypi.org/project/beautifulsoup4/.
[4] See https://komprehend.io/.
[5] See https://textblob.readthedocs.io/en/dev/quickstart.html.
[6] See https://pypi.org/project/vaderSentiment/.
[7] See https://pypi.org/project/text2emotion/.

there are several 'bandits' as an airport authority that usually collaborates with elected officials and the air traffic control regulator coordinates communication and messaging to community residents. Each agent learns from each other's actions whose probability of rewards will differ. Whereas 'action' and 'reward' are the two areas of the k multi-armed bandit, a contextual bandit framework includes 'state,' 'action,' and 'reward' as illustrated in Fig. 1.

The input parameters (theta or θ) in Fig. 1 involve the class of bandit (a Bernoulli bandit), the type of policy (we evaluate six algorithms described in the next section), the agent (combination of policy and bandit), simulations (1000 and a horizon of 30 periods), as well as the context (the number of arms and the number of features per arm). A Bernoulli bandit refers to an optimization problem where an agent sequentially pulls one of two arms. At each round, each bandit behaves like a random variable such that $Y_k \sim$ Bernoulli(θ_k).

The CMAB problem fits a community engagement environment because agents can compromise between exploiting known opportunities and exploring unknown ones while putting actions on pause to address uncertainty and ignorance about future actions. Van Emden and Kapstein (2018: 2) described the difference between CMAB and MAB in these terms: "In contextual bandit problems, CMAB policies differentiate themselves, by definition, from their MAB cousins in that they can make use of features that reflect the current state of the world—features that can then be mapped onto available arms or actions. This access to side information makes CMAB algorithms yet more relevant to many real-life decision problems than their MAB progenitors." The current state of the world is reflected in the sentiment of the community residents reflected in the sample.

4. Probability distributions and computation of cumulative regrets

We compared the outcomes of six algorithms using the *Contextual* package in R (Van Emden and Kaptein, 2018). We focused on cumulative regrets and the standard deviation of cumulative regrets as a basis of comparison. The computation involved 1000 simulations and a time horizon of 30.

Most models used in the literature of contextual bandits have been linear. There have been efforts to expand modeling to non-linear contextual bandits (Bubeck and Cesa-Bianchi, 2012; Valko et al., 2013) and neural-linear bandits (Zahavy and Mannor, 2019). In this study, we compare the outcomes of different linear models.

- *Oracle.* It serves as the baseline and indicates the regret probabilities when playing the optimal arm or selecting the optimal policy.
- *UCB1.* The upper confidence bound determines the estimated regret of each action and helps identify the action characterized by the lowest estimate. Auer et al. (2002) explained the algorithm in detail. According to Bilgin (2020:41), "Upper confidence bounds (UCB) is a simple yet effective solution to exploration-exploitation trade-off. The idea is that at each time step, we select the action that has the highest potential for reward. The potential of the action is calculated as the sum of the action value estimate and a measure of the uncertainty of this estimate. This sum is what we call the upper confidence bound." Based on Bilgin (2020), the formula to compute the UCB is as follows:

$$A_t \triangleq \underset{a}{\operatorname{argmax}} \left[Q_t(a) + c\sqrt{\frac{\ln t}{N_t(a)}} \right] \tag{5}$$

- *Thompson Sampling.* It uses a beta-binomial model with two parameters: alpha and beta. The outcomes generated from pulling arms in a previous step are sampled using the beta-binomial distribution and then the arm with the lowest value is selected if the goal is to minimize regrets. See Agrawal and Goyal (2012) for an exposition of Thompson sampling in MAB.
- *Epsilon Greedy.* Van Emden and Kaptein (2018: 29) explained that "Contrary to the previously introduced ε-first policy, an ε-greedy algorithm […] does not divide exploitation and exploration into two strictly separate phases—it explores with a probability of ε and exploits with a probability of $1 - \varepsilon$." Although ε-greedy action selection is a popular means of balancing exploration and exploitation, it chooses equally among all actions when it explores.
- *Softmax:* The selection of the arm follows a Boltzmann probability distribution. It uses tau (τ) as a parameter to determine how many arms can be explored. A high τ value implies that all arms are explored equally. Softmax improves on ε-greedy because it allows to vary the action probabilities as a graded function of the estimated value. While the greedy action is still given the highest selection probability, all the others are ranked and weighted according to their value estimates. These are called 'softmax action selection' rules. Below is the Boltzmann formula:

$$\frac{e^{Q_t(a)/\tau}}{\sum_{b=1}^{n} e^{Q_t(b)/\tau}} \tag{6}$$

As τ (also called the temperature parameter) tends to zero, the softmax action selection becomes the same as the greedy action.

- *Exp3*: The algorithm uses a mixture of uniform distribution to define a policy and assigns each action an exponential mass distribution to characterize the cumulative rewards.

The classification of sentiment into the three categories served to derive the expected regrets and rewards. The standard deviation provided a measure of volatility. Regret represents an important concept in CMAB because it measures the quality of an exploration algorithm. Regret can be defined as the difference between a payoff (a reward or return) from an action and the payoff from an action that has been implemented. According to Lonza (2019:289), "regret is defined as the opportunity lost in one step that is the regret, L, at time t, is as follows: $L_t = V^* - Q(a_t)$, where V^* is the optimal value and $Q(a_t)$, the action-value of a_t." As a result, when seeking to balance exploration and exploitation, we aim to minimize the cumulative regrets defined as

$$L_t = \Sigma_i (V^* - Q(a_i)) \tag{7}$$

Among the algorithms used in the simulations, we selected the UCB1 algorithm. According to Bilgin (2020:41), "an action is selected either because [the] estimate for the action value is high, or the action has not

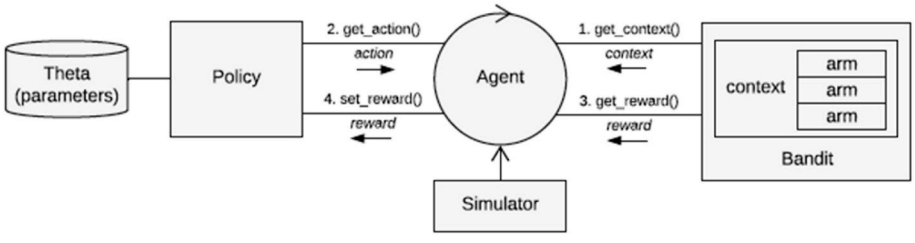

Fig. 1. Diagram of basic structure of the 'contextual' framework (Van Emden and Kaptein, 2018: 6).

been explored enough (i.e., as many times as the other ones) and there is high uncertainty about its value, or both."

8. **Determination of Policy Strategies.** The categories of sentiment and emotions served to predict policies in three categories: (1) 'exploit,' if the percentage of positive sentiment was higher than and the percentage of angry emotion was less than the respective average over the sampled period; (2) 'explore,' if the percentage of positive sentiment was lower than and the percentage of angry emotion was higher than the respective average over the same period; and (3) 'pause', otherwise.
9. **Classification Modeling.** In this analysis, 25 percent of the data for each sample year were set aside for the test dataset. We used a stacking ensemble model to optimally combine three machine learning classifier algorithms, namely, the K-Nearest Neighbors (KNN), Random Forest (RF), and Support Vector (SVC). The stacking model is a meta-learner algorithm (based on a Decision Tree Classifier) that combines the predictions from the three classifier algorithms to capitalize on the strengths of each algorithm and to improve the predictive accuracy of the overall model.

The process to build the stacking ensemble model consists of (1) splitting the training data into two folds; (2) choosing the weak learners (in the KNN, RF, and SVC models) and fitting them to the training data of the first fold; (3) making each weak learner predict observations in the second fold; and (4) fitting the Decision Tree Classifier meta-model on the second fold by leveraging the predictions made by the three algorithms as inputs. The split in the dataset into two folds is designed to make sure that data used for the training of the weak learners are not used for the meta-model. Below is a brief description of the weak learner algorithms. A more in-depth explanation of the algorithms is out of the scope of this article. Readers interested to stacked generalization are referred to Wolpert (1992).

The *K-Nearest Neighbor* is a non-parametric algorithm that calculates the distance between data points, find the closest neighbors, and vote for labels (class fit). Like Random Forest and Support Vector Machine, the K-Nearest Neighbor algorithm can be used for regression or classification. Fix and Hodges (1951) developed the model, which was later expanded by Cover and Hart (1967).

The *Random Forest* tree is an ensemble machine learning model that fits decision tree classifiers on sub-samples of the dataset. The power of ensemble models is that in the case of a small sample they can generate uncorrelated trees working in parallel through bootstrapping. These trees operate as an ensemble and the algorithm identifies the solution in the case of classification through voting. See Breiman (2001) for an explication of Random Forest models.

The Support *Vector Machine* algorithm constructs a hyperplane to identify classes based on support vectors. The SVC algorithm can determine the classes based on different kernels (linear, polynomial, and radial-based function). In our case, the optimal model based on a grid search model had a polynomial kernel, with C = 0.1 and gamma = 1. The C parameter determines the tolerance for misclassification or mismatch: the larger the C parameter value, the narrower the margin that separates classes, and the higher the model bias. The gamma parameter determines the outreach of the support vectors that define the margins. See Cortes and Vapnik (1995) for a description of SVM networks.

10. **Prediction of Policy Strategies.** Below are the tools used to measure the performance of the meta-learner model:

The first tool is the *classification report* that includes four key metrics including.

- Accuracy $= \frac{True\ Positive+True\ Negative}{True\ Positive+True\ Negative+False\ Positive\ +False\ Negative}$
- Precision $= \frac{True\ Positive}{True\ Positive+False\ Posiive}$
- Recall $= \frac{True\ Positive}{True\ Positive+False\ Negative}$
- $F-1\ \ score\ \ =\ \ 2\,x\ \frac{Precision\ x\ Recall}{Precision+Recall}$

If the model predicted 'exploration' and it was true, then we qualified the outcome as 'true positive.' The classification report provides metrics for two types of average. In the case of macro average, *each class* is weighted equally—whereas micro average refers to the case where *each sample* is weighted equally. Weighted average considers how many observations of each class there were in the computation of the classification metrics.

The second tool to evaluate the performance of a classification model is the *confusion matrix*, which is a summary of the predictions by class. The number of correct and incorrect predictions are used to compute true positive/true negative and false positive/false negative scores.

Finally, the Graphiz[8] library in Python shows the outputs of the meta-learner in the form of a tree including the Gini impurity index. An impurity of zero is the best impurity outcome, which is possible when all elements belong to the same class. If we have C total classes and p(i) is the probability of picking a datapoint with class i, then the Gini impurity index is calculated as

$$G \ = \sum_{i=1}^{C} p(i)^{*}(1-p(i)) \tag{8}$$

5. Analytical outcomes

5.1. Evaluation of sentiment, emotions, and intent

Table 1 provides the key sentiment analysis metrics by sampled year. Sentiment among communities fluctuated between 2015 and 2017 and turned more negative from 2018 to 2021.

The overall sentiment was positive in 2015. The data showed that residents expressed their appreciation for a cross-agency initiative involving ice rescue training exercises. They were thankful for the FAA's willingness to address operational changes to mitigate aircraft noise. Residents expressed enthusiasm because local elected officials were willing to get involved in abetting noise. The joint engagement of the FAA and the airport authority explained why a high degree of advocacy was detected in the content of the sample in 205 compared with subsequent years (Table 2).

Sentiment changed in 2016 when performance-based navigation (PBN) procedures were implemented. Satellite-guided PBN procedures allow equipped aircraft to follow a narrower path to and from an airport. PBN is designed to ensure on-time predictability and better access into congested metropolitan areas where large airports are closely located to smaller ones (mainly general aviation airports). According to collected data, some communities located on departure paths complained that they were increasingly subject to noise exposure because aircraft operations were concentrated on narrower flight paths.

The higher percentage of positive sentiment in 2017 reflected optimism from residents and local politicians about the possibility to work with the airport community and the FAA on addressing noise issues. Airport stakeholders were also supportive of the new air service. Nevertheless, a group of residents expressed opposition to early morning flight departures and increased concentration of flight paths over specific locations. The year 2017 represented a pivotal one in the sample because 57.8 percent of the measured intent represented feedback, which could be broken up into 6.7 percent expressing some form of appreciation, 65.5 percent voicing complaints, and 27.9 percent suggesting some improvement in abating noise.

[8] See https://graphviz.org.

Table 1
Key sentiment analysis metrics (percent).

	Sentiment			Subjectivity	Emotions				
Year	Positive	Neutral	Negative		angry	fear	happy	sad	surprise
2015	88	11	1	44	4	48	9	20	20
2016	5	24	71	43	5	44	8	25	18
2017	69	27	4	41	5	46	8	21	20
2018	22	35	43	43	4	47	9	20	20
2019	10	26	64	44	5	41	7	23	23
2020	13	30	57	43	4	48	11	22	15
2021	12	15	73	47	2	51	9	21	16

Table 2
Intent analysis (percent).

	Intent					
Year	Informational	Feedback	Query	Advocacy	Non-Informational	
2015	54.9	21.0	0.4	20.5	3.2	100
2016	14.3	57.8	9.6	3.1	15.2	100
2017	80.0	3.9	1.2	11.7	3.2	100
2018	65.4	9.4	0.2	11.5	13.5	100
2019	66.8	10.4	12.2	3.0	7.6	100
2020	37.8	19.5	24.2	4.0	14.5	100
2021	71.1	9.1	10.5	7.1	2.2	100

Sentiment turned more negative in 2018 and 2019. In 2018, residents expressed their concerns about the effects of noise-related cardiovascular diseases as well as the impact of particle pollution from aircraft on the incidence of cancer, pulmonary, and other diseases. The publication of medical research focusing on the impact of noise on health triggered a debate among community stakeholders. In 2018, the sample contained a high informational (65.4 percent) and a high non-informational (13.5 percent) content that denotes reports by health experts on the dangers of aircraft noise in digital prints and more spam in social media.

In 2019, digital prints showed that residents expressed frustration about perceived inaction on demanded changes in arrival and departure procedures. This translated into an increase in queries and questions about what the airport authority was doing to help mitigate noise.

Although the pandemic caused the number of operations to decline drastically in 2020, the airport community remained engaged as the percentage of feedback and query in the sampled content indicated (respectively, 19.5 and 24.2 percent respectively). A higher proportion of the content (11 percent) expressed happiness among airport community residents. Nevertheless, as more people teleworked, they were experiencing noise to which they were not usually exposed. This may explain the high proportion of fear (48 percent) and sadness (22 percent).

In the 2021 sample, the increase in negative sentiment analysis to 73 percent revealed that community residents were increasingly fearful about noise exposure due to growing operations. High negativity resulted from fear about the future and sadness that their situation will not change. This sentiment was prevalent in social media.

Subjectivity measures utterances on a scale from 0 percent (very objective) to 100 percent (very subjective). The metric indicates whether the content of information is more likely to reflect individual opinions than facts. The higher the subjectivity, the greater the emotional content. Table 1 shows that subjectivity has remained mostly stable throughout the sampled year until 2021.

The evolution of emotions reflected an erosion of support (except in 2017) among increasingly 'angry' residents, which was later supported by elected officials at the federal level. The proportion of 'gloomy' residents increased because they did not feel they had enough power to change navigational approaches and departure procedures. Sentiment and emotions are not sufficient to explain the relationships among stakeholders in a community. It is also necessary to evaluate the intent conveyed in the sample. Table 2 shows the categories of the intent underlying the sampled data.

The intent analysis shows whether sentiment is driven by facts (high informational content) or emotions (high non-informational content). It also categorizes the content that underlies whether there is any dialogue among stakeholders in the forms of feedback and queries or advocacy on specific issues. Following the deployment of new navigation procedures in 2016, residents around the airports complained about increased noise (57.8 percent feedback), which pushed airport management and regulators to explain the goals of the airspace redesign and the roles of the new procedures through workshops reported in the press and social media in 2017 (80 percent of sampled data were informational).

5.2. The computation of cumulative regrets and standard deviation

Table 3 compares the results of 1000 simulations over a time horizon of thirty episodes for six algorithms.

Bandit algorithms seek to minimize regret. Table 3 features the cumulative difference in the payoffs between possible and actual action over the simulations. In most cases, UCB1 provided the lowest standard deviation of cumulative regrets, which meant lower volatility.

6. Predicting the optimal policy

This section provides important metrics to help agents decide on a course of action.

Table 4 shows overall, the meta-learner had an accuracy of 86 percent. Individually, the KNN, the Random Forest, and the Support Vector classifier had an accuracy of respectively 57 percent, 85 percent, and 57 percent. Eighty percent of the predictions for 'exploration' are relevant and 100 percent of the total relevant prediction outcomes for 'pause' are correctly classified by the meta-learner. The weighted average is higher than the macro one because there were fewer of one class ('exploration') in the computation of the precision, recall, and F1-score. The macro average does not consider the proportion for each label in the dataset.

Based on the confusion matrix outputs in Table 5, we can determine that precision (considering weighted average) is equal to $\frac{6}{8} \approx 0.75$. This means that when the model predicts 'exploitation,' it is correct 75 percent of the time. Recall is equal to $\frac{6}{7} \approx 0.86$. This implies that the

Table 3
Cumulative regrets and cumulative regret standard deviation.

			2015		2016		2017	
			Cumulative	Cumulative	Cumulative	Cumulative	Cumulative	Cumulative
Algorithm	t	sims	Regrets	Regrets (sd)	Regrets	Regrets (sd)	Regrets	Regrets (sd)
Oracle	30	1000	0.00	0.00	19.82	6.81	0.00	0.00
UCB1	30	1000	6.78	1.27	6.31	3.27	6.10	3.69
Thompson Sampling	30	1000	3.67	2.40	4.34	4.88	4.24	4.90
Epsilon Greedy	30	1000	4.84	30.77	4.87	26.58	4.80	2561.00
Softmax	30	1000	2.61	14.00	3.55	20.01	3.65	20.15
Exp3	30	1000	14.89	11.21	11.19	10.00	9.75	10.29

			2018		2019		2020	
			Cumulative	Cumulative	Cumulative	Cumulative	Cumulative	Cumulative
Algorithm	t	sims	Regrets	Regrets (sd)	Regrets	Regrets (sd)	Regrets	Regrets (sd)
Oracle	30	1000	6.32	11.92	16.28	8.81	13.28	10.57
UCB1	30	1000	2.59	7.65	5.73	4.17	5.09	4.84
Thompson Sampling	30	1000	2.34	7.39	4.71	22.56	4.16	6.22
Epsilon Greedy	30	1000	2.25	9.64	3.69	18.37	4.19	16.96
Softmax	30	1000	2.12	8.34	3.69	18.37	3.51	14.98
Exp3	30	1000	2.82	10.68	9.19	9.66	7.09	9.95

			2021	
			Cumulative	Cumulative
Algorithm	t	sims	Regrets	Regrets (sd)
Oracle	30	1000	6.32	11.92
UCB1	30	1000	2.59	7.65
Thompson Sampling	30	1000	2.34	7.39
Epsilon Greedy	30	1000	2.25	9.64
Softmax	30	1000	2.12	8.34
Exp3	30	1000	2.82	10.68

Table 4
The classification report.

Policy	Precision	Recall	F1-Score
1: Explore	0.80	1.00	0.89
2: Pause	1.00	1.00	1.00
3: Exploit	0.00	0.00	0.00
Accuracy			0.86
Macro Average	0.60	0.67	0.63
Weighted Average	0.74	0.86	0.79

Table 5
Confusion matrix.

		Predicted		
		Explore	Pause	Exploit
Actual	Explore	0	0	1
	Pause	0	2	0
	Exploit	0	0	4

meta-learner identifies 86 percent of the time 'exploitation' when it is the selected policy.

We start in a position of 'exploitation' as the root node (depth zero). The Gini index is more than zero (G = 0.571) because the seven samples contained within the nodes belong to different classes. The values [4,2,1] indicate the number of samples, that is, seven. The value of '4' indicates the sample belongs to 'exploitation,' '2' to 'pause,' and '1' to 'exploration.' The class shows the predictions at a given node. If the predictions were to end at the root node, then it would predict that all seven samples belonged to the 'exploitation' class. If the strategy is 'pause,' then there are two samples and all of them are classified as 'pause,' hence the Gini impurity value of 0.

The Gini impurity index and entropy help derive the feature importance at each node. Entropy is a criterion for calculating information gain, which determines how a node is split. Among the features, the degree of 'anger' expressed by community residents is the most important one because it provides the information that an agent needs to choose a course of action. In the present case, an agent continues an 'exploitation' strategy until the level of 'anger' elicits a 'pause' strategy expressed in the matrix as $\pi = [0,1,0]$.

Finally, Table 6 shows a table of probability associated with each policy based on the stacking ensemble model. Over the seven intervals, the dominant strategy was 'exploitation' alternating with 'pause.' The limited choice of exploration may be explained by the fact that residents were vocal about fear of changes in their way of life and sadness due to their perceptions that politicians, the airport authority, and the regulator were not listening to their concerns.

7. Final remarks

The contextual multi-armed bandit problem made it possible to frame the environment of agents in community engagement and allowed the assessment of regrets related to the three categories of policy. In our study, we linked sentiment analysis with reinforcement learning through a stacking ensemble classification model designed to predict the incidence of sentiments and regrets on the classification of policy.

This study suggests that agents were more likely to exploit at first and then pause. Exploration did not appear to be a viable policy because agents could not stem both an increase in anger and disengagement as the sample analysis indicated. Agents must pay attention to 'angry'

Table 6
Stacking ensemble model predictions.

Time	Exploit	Pause	Explore
1	0.80	0.00	0.20
2	0.00	1.00	0.00
3	0.80	0.00	0.20
4	0.80	0.00	0.20
5	0.00	1.00	0.00
6	0.80	0.00	0.20
7	0.80	0.00	0.20

residents whose 'voice' may provide some insights on what needs to be changed in their course of action. For instance, in 2016, out of 57.8 percent of feedback provided in digital prints, 6.7 percent were appreciation, 65.5 percent were complaints and 27.9 percent were suggestions. Therefore, continuous monitoring of the content of media, sentiment, and emotion categories is necessary to monitor and evaluate engagement policy.

The case study is representative of airports where the implementation of new navigation procedures has resulted in more concentrated flight paths. In a context of strong opposition to the implementation of navigation procedures at a large hub airport, the classification models suggested that agents may not be better off by exploring new policies. The stacking ensemble model determined that the probability of predicting the selection of exploration when it was true was zero percent. Based on seven years of data, the model predicted that agents would start with exploitation as a dominant strategy and then select pause.

8. Conclusion

There are many examples of key projects at airports around the world that have been either delayed, sometimes for decades, or even canceled due to strong opposition from airport community residents. Agents can use their experience or listen to what people around them are expressing in various media.

While airlines use sentiment analysis extensively to improve passenger service, airports and government regulators have been lagging in leveraging the power of sentiment analysis to guide agents on how to engage communities. This research fills an important gap in the literature on community engagement. Sentiment analysis can inform agents and measure risk before they select a specific policy. It enables agents to be initiative-taking and deal with issues before they become problems. It provides an assessment of the environment, especially when relationships among stakeholders are volatile and/or antagonistic.

In this study, we proposed a policy framework to help agents select among policy alternatives. We demonstrated the significance of sentiment analysis in predicting the selection of a policy. The complexity of relationships among the community's stakeholders and the uncertainty of the environment both make it difficult to evaluate the risk of selecting one policy at the expense of another. Agents can 'exploit' the benefits from implemented policies, 'explore' new policy directions, or 'pause' to re-assess the environment before acting.

Agents can try different policies to minimize negative payoffs. They can capitalize on their experience or read the polls. Eventually, they must operationalize sentiment and emotion to minimize risks and respond appropriately to environmental uncertainty. Examples around the world have shown that local politicians, airport authorities, and regulators have not always been successful to overcome popular discontent linked to airport expansion or noise. This article provides a methodology that airport community stakeholders can use to inform a course of action and quantify costly trial and error that may result in costly project delays or cancellations.

The selection of 'pause' or policy re-evaluation should not be perceived as inaction or lack of decision-making on behalf of agents. The 'pause' policy can help answer two questions: "Did the policy produce its intended outcome?" "Of what value is the outcome?" The worth of a policy is in residents' perception of the policy benefits, which can be translated into sentiment. Agents should explain the purpose of 'pause' to their audience, so it is not misconstrued as inaction. This leaves the door open to further research on how to identify the alternate optimal policy. There are other modeling perspectives such as Temporal Dependency that can be contrasted to the CMAB model presented in this paper.

Data availability

Data will be made available on request.

Acknowledgement

This research was conducted independently, and it does not reflect the opinion of the Federal Aviation Administration.

References

Agrawal, S., Goyal, N., 2012. June). Analysis of Thompson sampling for the multi-armed bandit problem. In: Conference on Learning Theory. JMLR Workshop and Conference Proceedings (23), 39-1 – 39-26.
Auer, P., Cesa-Bianchi, N., Fischer, P., 2002. Finite-time analysis of the multiarmed bandit problem. Mach. Learn. 47 (2), 235–256.
Bateman, I.J., Carson, R.T., Day, B., Hanemann, M., Hanley, N., Hett, T., et al., 2002. Economic Valuation with Stated Preference Techniques: a Manual. *Economic Valuation with Stated Preference Techniques: a Manual.* Edward Elgar publishing.
Baumol, W.J., Oates, W.E., 1988. Theory of Environmental Policy, second ed. Cambridge university press.
Bilgin, E., 2020. Mastering Reinforcement Learning with Python. Packt publishing.
Bouneffouf, D., Rish, I., 2019. A Survey on Practical Applications of Multi-Armed and Contextual Bandits. *arXiv preprint arXiv:1904.10040*.
Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5–32.
Bubeck, S., Cesa-Bianchi, N., 2012. Regret Analysis of Stochastic and Nonstochastic Multi-Armed Bandit Problems. *arXiv preprint arXiv*:1204.5721.
Chapelle, O., Manavoglu, E., Rosales, R., 2014. Simple and scalable response prediction for display advertising. ACM Trans. Intell. Syst. Sci. Technol. (TIST) 5 (4), 1–34.
Ciaburro, G., 2019. Hands-on Reinforcement Learning with R. Packt publishing.
Coase, R.H., 1960. The problem of social cost. J. Law Econ. 3, 1–44.
Cortes, C., Vapnik, V., 1995. Support vector networks. Mach. Learn. 20 (3), 273–297.
Cover, T., Hart, P., 1967. Nearest neighbor pattern classification. IEEE Trans. Inf. Theor. 13 (1), 21–27.
Diana, T., 2021. Improving messaging to airport community residents: an application of sentiment analysis to community engagement. J. Airpt. Manag. 15 (3), 304–312.
Dunn, W.N., 2015. Public Policy Analysis. Routledge.
Fix, E., Hodges, J., 1951. Discriminatory Analysis: Nonparametric Discriminatory Consistency Properties. U.S. school of aviation medicine, Randolph Field, *Texas*, February.
Hirschman, A.O., 1970. Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States, vol. 25. Harvard university press.
Lonza, A., 2019. Reinforcement Learning Algorithms with Python. Packt Publishing.
Palmas, A., Ghelfi, E., Galina Petre, A., Kulkarni, M., Anand, N.S., Nguyen, Q., Sen, A., So, A., Basak, S., 2020. The Reinforcement Learning Workshop: Learn How to Apply Cutting-Edge Reinforcement Learning Algorithms to a Wide Range of Control Problems. Packt publishing.
Robbins, H., 1952. Some aspects of the sequential design of experiments. Bull. Am. Math. Soc. 55, 527–535.
Salvador, J., Oliveira, J., Breternitz, M., 2020. Reinforcement Learning: a Literature Review. https://doi.org/10.13140/RG.2.2.30323.76327 retrieved at. https://www.researchgate.net/publication/344930010_REINFORCEMENT_LEARNING_A_LITERATURE_REVIEW_September_2020?channel=doi&linkId=5f9975b8a6fdccfd7b84d84a&showFulltext=true. (Accessed September 2020).
Sutton, R.S., Barto, A.G., 2018. Reinforcement Learning: an Introduction. MIT press.
Valko, M., Korda, N., Munos, R., Flaounas, I., Cristianini, N., 2013. Finite-time Analysis of Kernelised Contextual Bandits. *arXiv preprint arXiv*:1309.6869.
Van Emden, R., Kaptein, M., 2018. Contextual: Evaluating Contextual Multi-Armed Bandit Problems in R. *arXiv preprint arXiv*:1811.01926.
Villar, S.S., Bowden, J., Wason, J., 2015. Multi-armed bandit models for the optimal design of clinical trials: benefits and challenges. Stat. Sci. 30 (2), 199–215.
White, J.M., 2012. Bandit Algorithms for Website Optimization. O'Reilly.
Winder, P., 2021. Reinforcement Learning: Industrial Applications of Intelligent Agents. O'Reilly.
Wolpert, David H., 1992. Stacked generalization. Neural Network. 5 (2), 241–259.
Zahavy, T., Mannor, S., 2019. Deep Neural Linear Bandits: Overcoming Catastrophic Forgetting through Likelihood Matching. *arXiv preprint* arXiv:1901.08612.