Explaining Active Learning Queries

Kyu Hyun Chang
Advisor: Sorelle Friedler

May 2, 2017

Abstract

In contrast to traditional supervised machine learning that takes a set of labeled data and builds a model that best fits the given data, active learning selects instances from which it will learn. In a typical setting, active learning starts with some labeled instance, and queries an unlabeled instance that it can learn the most from. Then the queried instance is labeled by an oracle, and the learner re-trains the model and continues the learning cycle. By selecting the most informative instances, active learning attempts to find the optimal set of training data.

Often, the oracle is a human annotator: for speech data, for example, an annotator can be a trained linguist. In a typical active learning setting, an annotator's role is to provide a label to the instance that the active learner asks for. In this setting, it is difficult for the annotator to understand why the queried instance is important, and the annotator takes a passive role in a sense that he or she merely provides the label to the active learner. In this paper, I propose a technique that explains active learning queries and an expert-aided active learning procedure in which experts are more involved in the learning cycle.

The technique was applied to Haverford's Dark Reactions Project dataset, which consists of organically-templated metal oxide synthesis reactions. The explanations of queries are provided to a chemist, who was able to interpret the explanations and found it helpful identifying chemical space that is poorly understood by the model. Moreover, the expert-aided active learning showed performance commensurate with the standard active learning.

Contents

1 Introduction

1.1 What is Active Learning?

Traditional supervised machine learning takes a set of labeled data and builds a model that best fits the given data. This approach works well when there are a plenty of data to be trained from. However, in some cases, labeling can be time-consuming or expensive, and thus only a limited number of labeled data items can be obtained. Consider speech data; annotating these data is time-consuming and requires trained linguists.

In contrast to traditional supervised machine learning, *active learning* (also called *optimal experimental design* in the statistics literature) takes a different approach regarding training data. Rather than simply training a model with a given set of training data, active learning selects data from which it thinks it will learn the most. By selecting data (Section 1.3) from which it will learn, active learning attempts to achieve higher accuracy than comparable supervised learning with the same amount of labeled data.

To demonstrate how active learning can outperform traditional supervised learning with random sampling (*passive learning*) [29], I generated the following toy example. In this example, 200 instances are sampled evenly from two normal distributions. Then logistic regression is used for a classification task. Figure 1 and Figure 2 show the results for active learning and passive learning respectively.

Active learning used uncertainty sampling (Section 1.3.1), in which the learner queries the instance that the model is the least confident about. With 20 labeled instances, active learning achieved 97% accuracy (3% error), while the random sampling showed 85% accuracy (15% error). For this dataset, the instances that are close to the boundary are more informative than the ones that are far from the boundary. Active learning was able to outperform passive learning by successfully querying the instances that the model can learn the most from.

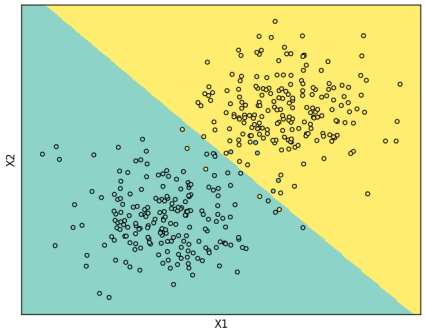

Figure 1: Uncertainty Sampling

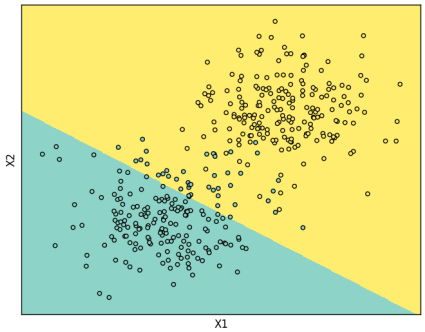

Figure 2: Random Sampling

1.1.1 Rationales For Active Learning

There are two rationales for how active learning can be helpful. One rationale is *exploitation of cluster structure* [10]. Suppose there are three clusters and three classes in data as shown in Figure 3. In this example, it is sufficient to have three labeled instances to correctly label the rest of the instances. For example, a Support Vector Machine (SVM), which classifies based on hyperplanes that separate instances of different classes with the largest gap as possible, can effectively classify the rest of instances if the labels of the three circled instances are given. In Figure 3, consider the lines as the decision boundaries or hyperplanes. Active learning exploits cluster structure in data by specifically asking for the labels of the instances that the model is uncertain about or instances that are sufficiently different from the instances that are already in the training set. In reality, the problem is much more complicated: there might be no cluster in data or even if there are clusters, most of the instances in a single cluster should be the same class.

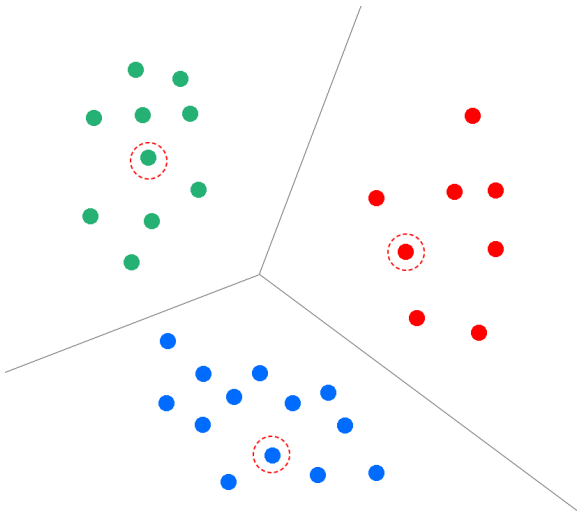

Figure 3: Clusters

The second rationale is an *efficient search through the version space*. The version space is a set of hypotheses consistent with the currently labeled data [24]. Whenever a newly labeled instance is added, the version space shrinks, and active learning selects an unlabeled instance that will shrink the version space the most.

In Figure 4 and 5, the red and blue points represent labeled instances where a color represents a class. The gray points represent unlabeled instances, and each rectangle represents a version in the version space. If an instance is within the box, the model classifies the instance as class A, otherwise class B. In Figure 4, all the boxes are valid versions and correctly classify all the labeled instances. However, when the circled instance

is labeled as A, two of the five versions in Figure 4 are no longer valid since they will misclassify the instance as class B. Thus, as depicted in Figure 5 the version space shrinks. Active learners query instances like this in search for the best version. The *Query-by-Committee algorithm* (Section 1.3.2) explicitly uses this idea by employing multiple models and querying the instances that the models disagree about the most.

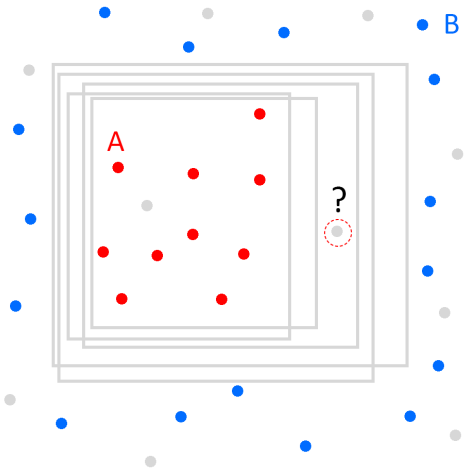

Figure 4: Version Space: Before Query

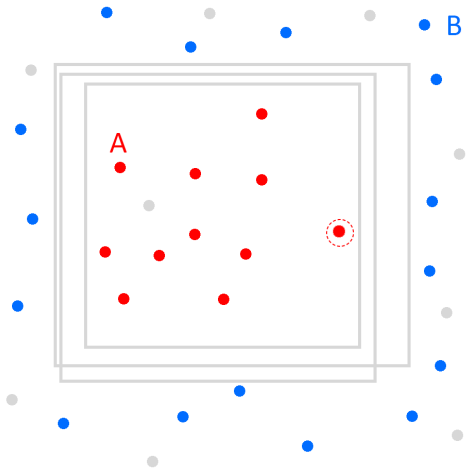

Figure 5: Version Space: After Query

1.1.2 Active Learning Setup

A general setup for active learning is where an active learner starts with a set of some labeled data or starts by randomly querying labels of some data. Then it trains its underlying machine learning model (e.g. SVM or artificial neural network) with the labeled data. Once the active learner trains its model, the learner queries label(s) of unlabeled data based on its selection strategy (Section 1.3). The labels are then provided by an oracle (e.g. a human annotator). The model learns from the newly labeled data and continues the process until it achieves a given accuracy threshold or until no more labeling is available (e.g. reached a budget constraint). Figure 6 shows an active learning cycle.

1.1.3 Applications of Active Learning

As briefly mentioned earlier, active learning is especially useful and effective when gaining unlabeled data is inexpensive or free, but labeling data is expensive or time-consuming. Some examples where active learning can be effective include:

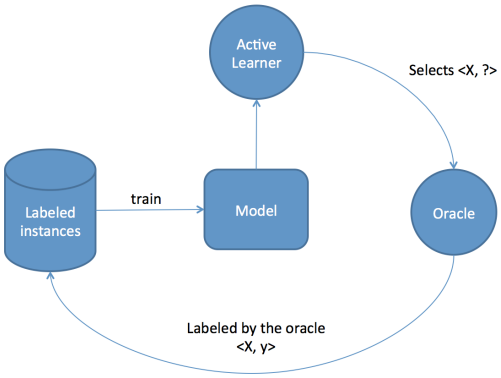

Figure 6: Active Learning Cycle

- Image Classification: For image classification, it is easy to obtain ample unlabeled images, but labeling images must be done by human annotators.
- Text Categorization: Consider a webpage categorization problem. Getting millions of webpages can be easily done by a web crawler, but annotating the categories of webpages is much more time-consuming and costly [14].
- Speech Recognition: Labeling speech data is time-consuming and often requires trained linguists. For example, transcription of speech utterance at word level takes ten times longer than the utterance duration (i.e. 10 minutes for a 1-minute length of audio). Even worse, transcription at phonetic level takes four hundred times longer than the utterance duration [39].
- Information Extraction: Information extraction often requires detailed labeling of a training set. The situation gets worse if labeling requires domain knowledge. For example, annotating genetic data and disease mentions of medical information often requires PhD-level biologists [29].

When labeling data is inexpensive, traditional supervised learning suffices. In reality, however, this is not always the case. When labeling data is expensive, achieving the same accuracy by using fewer labeled data items is directly related to the cost. Therefore, active learning is well motivated in many problems where obtaining unlabeled data is inexpensive, but obtaining labeled data is difficult, time-consuming or expensive.

1.2 Active Learning Settings

There are several active learning settings where a learner can query for a label of data. The most widely used ones are stream-based sampling, pool-based sampling, and membership query synthesis.

1.2.1 Stream-based Selective Sampling

Selective sampling (also called *stream-based* or *sequential* active learning) assumes that obtaining an unlabeled instance is free or inexpensive [2, 29]. In stream-based selective sampling, a potential query (an unlabeled instance) is drawn from the actual distribution, and the learner decides whether to query the instance or not. Such decisions can be based on a selection strategy (Section 1.3) with a certain threshold. Another approach is *region of uncertainty* [2]. In this approach, the active learner maintains certain regions where the underlying model is uncertain, and only queries the unlabeled instances that lie in these regions.

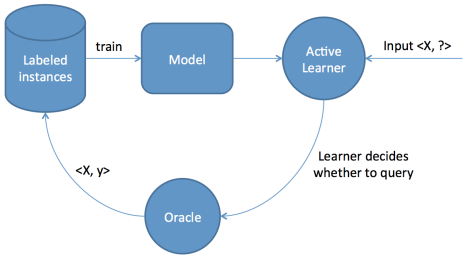

Figure 7: Stream-based Selective Sampling

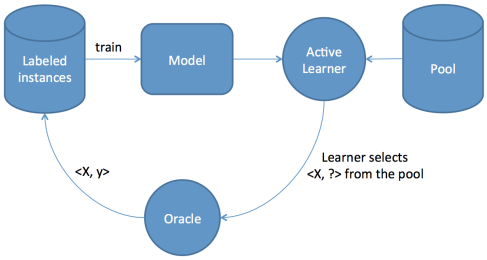

Figure 8: Pool-based Sampling

1.2.2 Pool-based Sampling

When obtaining a large set of unlabeled data is free or inexpensive *pool-based sampling* [19] is appropriate. For example, in computational chemistry, a large set of potential chemical experiments can be generated all together. Pool-based sampling assumes that there is a small set of labeled instances and a large set of unlabeled instances. The active learner evaluates the informativeness of each instance in the pool based on its selection strategy (Section 1.3). The most informative query is then labeled by the oracle and the learning cycle continues. Compared to stream-based selective sampling, pool-based sampling requires more computing power and memory. If there is a significant restriction on computing power or memory stream-based selective sampling could be more appropriate. Alternatively, instead of scanning the entire pool, a subset of the pool can be scanned when the pool size is too large. Since pool-based sampling selects queries that will give the most information, the set of training data collected based on pool-based sampling can be significantly biased (Section 2.1).

In both pool-based sampling and stream-based selective sampling, all queries are drawn from the actual distribution so the queries are guaranteed to be sensible. In membership

query synthesis, however, that is not always the case (Section 1.2.3).

Pool-based Sampling has been studied in many real-world problem domains such as speech recognition, text classification, information extraction, and image classification, and it has been the most popular scenario in applied active learning.

1.2.3 Membership Query Synthesis

Unlike pool-based or stream-based selective sampling, *membership query synthesis* [1] can select any unlabeled instance from the input space. In this setting, the active learner generates a synthetic instance rather than selecting among a pool of samples.

Since membership query synthesis has no restriction on what instances can be queried, a study shows that employing membership query synthesis reduces the predictive error rate more quickly than the comparable pool-based sampling [20].

On the other hand, in cases where the oracle is a human annotator, membership query synthesis can generate queries that are not sensible. An empirical study shows when membership query synthesis was employed for training a neural network to classify handwritten digits, the queries generated by the active learner were not sensible (e.g. an artificial hybrid word) to humans and had no natural semantic meanings [5].

Membership query synthesis has been shown to be effective for the domains where the oracle is not a human annotator. King et al. devised a "robot scientist" which carries a series of biological experiments for determination of gene function using mutants of yeast *Saccharomyces cerevisiae* and growth media [17]. In this setup, an instance consists of a growth medium (a mixture of chemical solutions) and a particular mutant of yeast. The "robot scientist" using membership query synthesis selects what experiments (unlabeled instances) to be conducted next. Then, the selected experiments are actually conducted to obtain the labels. The result shows the experiment selection strategy is competitive with human performance, and significantly outperforms a naive strategy (a strategy of choosing the cheapest experiment that has not been done) and random selection, showing a 3-fold and 100-fold decrease in cost respectively.

One can image that membership query synthesis can also be effective for setups such as drug discovery and materials discovery.

1.2.4 Query in Batches

So far, all query selection strategies assume that one query is being selected at a time. In practice, this can be highly inefficient. For instance, conducting biological or chemical experiments is time-consuming, but multiple experiments can be conducted at the same time. So it is more efficient to get a batch of queries and conduct several experiments and feedback a batch of labels (experiment results). In these setups, a batch-selection, where the learner queries a batch of instances rather than a single instance, is more suitable. For example, when active learning is used to assist a drug discovery process, a batch selection

is employed [36].

The most naive approach is *Q-best* where an active learner selects Q best instances. This approach, however, does not consider the information overlap among the query instances, and therefore it is not effective in most cases. In order to avoid this problem, there have been other approaches proposed such as incorporating diversity among the instances in batch and treating the selection process as a set optimization problem in which active learner selects a set of certain size that provides the most information [6, 13]

1.3 Query Selection Strategies for Active Learning

Active learning attempts to find the optimal training set by selecting a series of unlabeled instances that are most informative. There are several query selection strategies or learning criteria that have been studied in the active learning literature.

1.3.1 Uncertainty Sampling

Uncertainty sampling is a selection strategy where an active learner selects the least certain instance. For example, in a binary classification problem, the active learner will select a data point the current underlying model is least certain about (i.e. the one close to 0.5 confidence). This strategy is simple, but it only considers a single data point where the model is the least certain and does not make use of other labeled points [29]. Mathematically, the *least confident* instance, x^*_{LC}, is the following:

$$x^*_{LC} = \arg\max_x (1 - P(\hat{y}|x))$$

where $P(\hat{y}|x)$ is the probability (confidence) of the instance x having predicted outcome $\hat{y}$ under the current model.

There are other variants of the least certain selection strategy that use different uncertainty measures. *Margin sampling* makes use of the most and the second most probable class labels for its uncertainty measure [28, 29]. By incorporating the second most probable class, it attempts to make the best use of the distribution of labeled data.

In the following sections, x^* is the most uncertain instance, y is a label, and $P(\hat{y}|x)$ is the probability of instance x being $\hat{y}$ under the current model. Under *margin sampling*, x^*_M is the following:

$$x^*_M = \arg\min_x (P(\hat{y_1}|x) - P(\hat{y_2}|x))$$

where y_1 is the most probable and y_2 is the second most probable class of x. x^*_M is the instance that the current model is the most confused about meaning the model is not sure whether to classify x as y_1 or y_2.

Another variant of uncertainty sampling, perhaps the most widely used one, is Shannon *entropy* [33]:

$$x_H^* = \arg\max_x -(\sum_i P(\hat{y}_i|x) \log P(\hat{y}_i|x))$$

where $-\sum_i P(\hat{y}_i|x) \log P(\hat{y}_i|x)$ is the Shannon entropy, H. It can be interpreted as selecting an instance that will give the most information when the model knows about the label of the instance.

For binary classification problems, all three methods are mathematically equivalent.

1.3.2 Query by Committee Algorithm

The *Query-by-committee-algorithm* (QBC) [32] uses an *ensemble method* [12] along with uncertainty sampling. As the name suggests, the QBC algorithm maintains a "committee", $C = \{c_1, c_2, ..., c_n\}$, of machine learning models trained with the same labeled instance. In the QBC algorithm, each model, c_i, can be thought of as an explicit version or hypothesis for the current set of labeled instances. For a query selection, each model votes on labeling of unlabeled instances, and the instance where committee members disagree the most is selected as the next query.

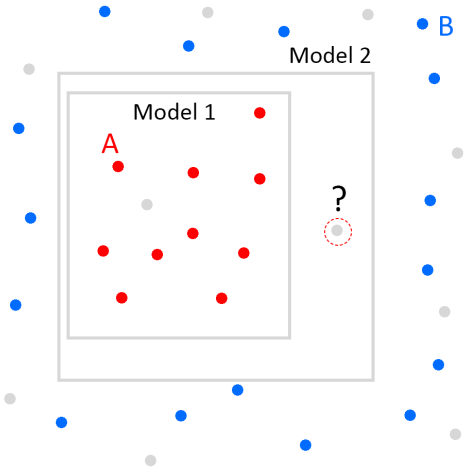

Figure 9: QBC - Before Query

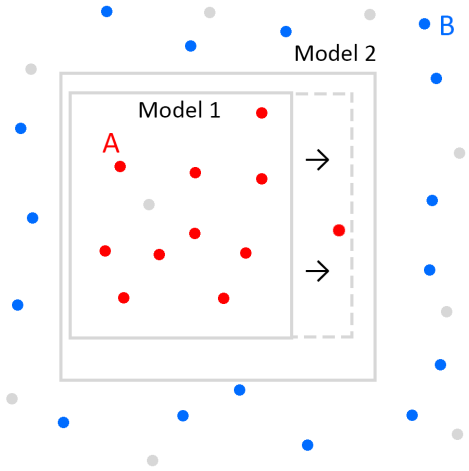

Figure 10: QBC - After Query

Figure 9 and 10 provides a visualization for the QBC algorithm. In these figures, the committee consists of two members: model 1 and model 2 are represented as boxes. Two members agree on all the instances except for the circled instance in 9. Since this instance is the instance where the committee disagrees the most, the active learner will query this instance. Once the instance is labeled, the version of model 1 needs to expand to correctly

label the newly labeled instance as shown in Figure 10. As active learning proceeds, the versions of members converges.

The intuition behind the QBC algorithm is minimizing version space. By quickly identifying instances that different models disagree with, it selects a set of labels where several models can converge, narrowing down version space.

1.3.3 Other Selection Strategies

There are several other selection strategies such as *expected error reduction* [8], *variance reduction*, and *expected model change [31]*:

- Expected Error Reduction: the learner queries the instance that would reduce the generalization error the most.
- Expected Variance Reduction: the learner queries the instance that would minimize the output variance.
- Expected Model Change, the learner selects the unlabeled instance that would change the current model the most if the label is known.

1.4 Sampling Bias in Active Learning

Although many empirical studies and experiments report that active learning outperforms comparable passive learning, there have been reports that active learning does not provide any benefit. In some cases, it was reported that active learning does worse than its comparable passive learning [10, 14].

The underlying reason why active learning under-performs is *sampling bias*. Sampling bias is a bias in which a sample is collected in a way that some instances are less likely to be included in the sample. In machine learning as well as in the statistics a training dataset or sample is assumed to represent the population. Since the training set is a representation of the population, the model is derived from the training set, and the model is expected to represent the population. If the training set is not an accurate representation of the intended population, then the erroneous model is likely to be derived. For example, let's say that one is trying to make a machine learning model for predicting a house price in New York City on the basis of some characteristics of a house. Suppose that the most of the training data is from the Bronx, and it lacks data from Manhattan. If a model is trained with this dataset, then it is likely that the model would fail to predict a price of a house in Manhattan.

Dasgupta and Hsu presents a one-dimension example, where an active learner underperforms a comparable model with random selection even with infinitely many labels [11].

In high dimensions, the problem arising from sampling bias is often worse. Many studies focused on improving uncertainty measures or learning criteria that reduce sampling bias [11, 14, 16, 38].

Due to its nature of selective sampling, the distribution of training set collected by active learning is tied to the model and the selection strategy. Thus, it is inevitable that it quickly diverges from the natural distribution.

Some studies report that training set collected with an active learner can still be effective in training a different model. Lewis and Catlett trained a decision tree (C4.5) model with the training set collected by an active learner with naive Bayes model and uncertainty sampling [18]. Although the training set is biased and selected by a different model, the decision tree trained with the training set yielded higher accuracy than the decision tree trained with a randomly selected training set that is ten times larger. Other studies also successfully re-used the training set collected by an active learner of a model to train different models [15, 34].

However, there are also opposite results. In Baldridge and Osborne's study, when a training set of a parser model is used to train different parser models, often the gains from active learning is negligible or sometimes even worse than random selection [3]. Baldridge and Osborne reduced the number of training instances required by each parser model by employing an active learner with a heterogeneous ensemble model that consists of several parser models. Intuitively, an ensemble model is a mixture of models thus a training set collected with an active learner that uses an ensemble model as an underlying model is expected to be more general than a training set collected with a single model active learner.

The main reason for using active learning is because obtaining labels of data is expensive. As Baldridge and Osborne's study shows a training set built via active learning might not be useful to different models [3], meaning one might need to discard the training set in order to avoid training a model from a biased sample. In cases that there is no known model or model needs to be changed, using active learning might result in an increase in total cost of labeling.

Although using heterogeneous ensemble method helps in generating a training set that is more applicable to different models, ensemble methods require more computing power than single model methods. Thus, an ensemble method may not be always feasible.

1.5 Further Readings

The background for the literature review is heavily borrowed from Settles's *Active Learning Literature Survey* [29] and *Active Learning* book [30]. The survey paper and the book provide a comprehensive background for active learning. Hanneke's *Theory of Active Learning* [14] also provides a good theoretical background of active learning with an emphasis on disagreement-based active learning methods.

2 Statement of the Problem

2.1 Explaining the Predictions of Any Classifier

Although machine learning has been widely adopted, understanding the reasoning behind machine learning models has been often neglected. For understanding an obscure model globally, the most common approach has been building a human interpretable model such as a decision tree based on the predictions of the obscure model such as SVM or artificial neural network [4, 9].

Analyzing the decision tree extracted from an underlying model provides a holistic understanding of the model and it helps to understand hypotheses of the underlying model. For example, Raccuglia et al. generated chemical hypotheses to guide future chemical experiments by using an SVM-derived decision tree [25].

Extracting and analyzing the decision tree derived from an obscure model provides an explanation of a model. However, using the decision tree to explain individual predictions may not be very useful. Explaining a prediction not only involves tracking down the decision tree from the root to a node but also tracking down the other branches to examine how one decision on a feature affects the prediction. For example, if a decision node is $X_1 < 0$ for a feature X then it is difficult to say how the value of X_1 affects the prediction without examining both branches and their sub-branches from the node unless one side of the branch always results in one class. Suppose an instance has a positive value for X_1. Depending on the values of the other features, having a positive value for X_1 may affect the instance to be classified as one class or another. It is also possible the value of X_1 does not affect the prediction at all given the values of the instance's other features. To examine how X_1 affects the prediction examinations of sub-branches may be required.

With decision trees, especially with the ones that involve many features, it is not an easy task to analyze and pinpoint what features were important and how those features affected the prediction of an instance.

2.1.1 Local Interpretable Model-agnostic Explanations

As discussed above, it has been a challenge to provide an interpretable explanation of a prediction. Recently proposed *Local Interpretable Model-agnostic Explanations* (LIME) [27] has demonstrated success in providing explanations for predictions. LIME is a model-agnostic technique to provide explanations for the predictions of a classifier. It is similar to methods used to explain a model in that LIME also derives interpretable models such as decision tree, linear models, and falling rule lists to provide an explanation. The key difference is that LIME builds an interpretable model with a generated sample in the vicinity of the prediction that it tries to explain, rather than building a global model.

For a given predicted instance, x, LIME generates a perturbed sample set $\mathcal{Z}$ in the neighborhood of the instance. Then based on the sample $\mathcal{Z}$ and their labels from the

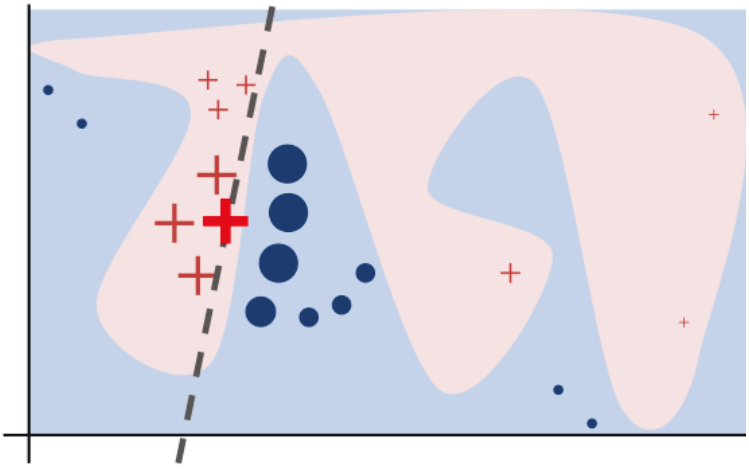

Figure 11: LIME [27]
The blue and red background represent the decision boundary of a complex model. For a prediction of an instance x, the bold red cross, that needs to be explained, LIME generates a sample in the vicinity of x by perturbing x. Based on the sample, LIME trains an interpretable model from which it will derive an explanation from (a linear model for this figure). Although this model does not necessarily explain the model globally, it approximates how the model behaves in the vicinity of x [26].

model, LIME searches for the most interpretable model, and derives an explanation, $\xi(x)$. The explanation is obtained by the following:

$$\xi(x) = \arg\min_{g \in G} \mathcal{L}(f, g, \pi_x) \; + \; \Omega(g)$$

where G is a set of potentially interpretable models such as linear models, decision trees, and falling rule lists. f is the complicated model whose prediction that LIME attempts to explain. $\Omega(g)$ is the complexity of an interpretable model g, which can be the number of variables for a linear model or the depth of a decision tree. $\mathcal{L}(f, g, \pi_x)$ is a measure of how unfaithful a model g or how inaccurate the model is in the vicinity of x. The unfaithfulness of a model is computed as the following:

$$\mathcal{L}(f, g, \pi_x) = \sum_{z,\ z' \in \mathcal{Z}} \pi_x(z)(f(z) - g(z'))^2$$

in which π_x is a weight (e.g. a distance measure between x and z').

Understanding the reasoning behind a model and especially understanding predictions are useful in several ways. An explanation of prediction is important in accessing trust of

a model [27]. Even if a model has shown to be highly accurate in a test set, the accuracy itself may not be enough for a human to determine to use the model for decision making.

Consider a machine learning model that aids doctors by providing a diagnosis of a patient. In order for a doctor to trust the prediction of the model, the reasoning behind the prediction is crucial. Without any explanation, a doctor can be reluctant to trust the model, but if the model provides an explanation such as "sneeze", "fever", and "inflammation" the doctor will be more willing to trust the model. Moreover, if the model provides an explanation that does not make sense based on the doctor's prior knowledge (e.g. "height" for cold), the doctor can easily access the appropriateness of the model. The explanation for prediction is not only useful to access the trust but also helpful to easily identify the weaknesses and flaws of a model using humans' prior knowledge and avoid a hasty conclusion of a model.

2.2 Problem Statement: Explaining Active Learning Queries

One problem in active learning is that when an active learner queries an instance, it does not provide a very human-interpretable explanation about the query. In the settings where annotators are human experts, it is natural that the experts ask "why this particular instance?"

As discussed in Query Selection Strategy for Active Learning (Section 1.3), a query is selected based on a metric, and the value of "uncertainty" whether it is least confident, margin, or entropy does not convey much information about the query instance. For example, the value of uncertainty measure such as entropy is not helpful in providing a human-interpretable explanation of "why" or "how" labeling the instance is likely to improve the model.

If a query is provided with an explanation of how or why labeling the query will improve the model, such information can offer insight to domain experts. The simplest form of such explanation would be to explain the key variables in the query that make it difficult to predict the outcome.

Active learning showed a promising result in computational chemistry. In a discovery process of finding active compounds from a large dataset of compounds, using an active learner outperformed using a passive learner with random sampling [35, 36]. In addition to simply providing a batch of experiments that are expected to improve the underlying model the most, if the active learner provides clues to why it thinks these experiments are important or how these experiments are going to improve the current model, such explanations of queries can offer insight to chemists.

Rather than blindly following the recommendations of an active learner, the cycle of active learning can be more interactive between an active learner and chemists (human annotators). For example, when explanations are provided along with queries, a hybrid active learning setting is possible, in which chemists use their domain knowledge to decide not to perform some recommended reactions and instead perform others that are more

intriguing based on the explanations of the queries.

In a traditional active learning setting, human annotators do not play an active role other than providing the labels of queries. When human annotators are experts, their knowledge is not utilized and the interaction between an active learner and domain experts is one directional. When a query is provided with an explanation, the process of querying and labeling can be more interactive and allow experts greater leeway to utilize their domain knowledge in a more interesting way. Moreover, incorporating domain knowledge in active learning may lead to faster learning.

3 Methods

In this section, I propose a technique that generates a human interpretable explanation of active learning queries using LIME. Although the method focuses on active learning of a binary classification problem of *uncertainty sampling*, it can be generalized and used with a different selection strategy or/and a multi-class classification problem. More discussion on applying the method to a different strategy is provided in Section 3.3.

3.1 Using LIME to Explain Uncertainty

As introduced in Section 2.1.1, LIME provides an explanation of a prediction. Directly applying LIME on a query instance provides an explanation of the prediction of the model, which is different from explaining why the query instance is selected.

In active learning with *uncertainty sampling*, a query instance is the most uncertain instance in the pool of unlabeled instances. Thus, explaining why the query instance is selected is reduced to explaining the uncertainty of the query instance.

3.1.1 Definition of Certainty

Given an instance, the certainty of the prediction of a model measures how confident the model is in predicting the class of the instance.

For multi-class classification with n classes, the certainty of a prediction of instance x, x^*_C, can be defined as follows:

$$x^*_C = \frac{n}{n-1}\left|P(Class_{highest\_prob}|x) - \frac{1}{n}\right|$$

where $Class_{highest\_prob}$ is the class with the highest probability and $P(Class_{highest\_prob}|x)$ is the probability of x being the class. The certainty value ranges from 0 (most uncertain) to 1 (most certain).

For binary classification, since $P(Class_0|x) + P(Class_1) = 1$, the certainty of a prediction of instance x, x^*_C, can be simplified as follows:

$$x^*_C = 2|P(Class_0|x) - 0.5|$$

where $P(Class_0|x)$ is the probability of instance x being $Class_0$ under the current model. The rest of the paper focuses on binary classification problems.

3.1.2 Uncertainty Mapping

For most of the widely used models such as SVM, Random Forest, and Neural Network, there are standardized methods to estimate probabilities of their predictions[1]. Using this estimated probability, a certainty value can be computed for each instance, and based on a threshold, instances are labeled as either certain or uncertain. While such a threshold can be a fixed value, it can also be dynamically determined. For example, the threshold can be set to the 25th percentile of certainty values computed for all unlabeled instances.

Figures 12 and 13 illustrate how *uncertainty mapping* works. Figure 12 shows a plot for a binary classification problem, where red and green points represent labeled instances and blue points represent unlabeled instances. For the base model, a SVM with a linear kernel is used.

Based on the trained SVM, all instances are mapped as either certain or uncertain as illustrated in Figure 13[2]. For this example, the threshold is dynamically determined based on the certainty values of the 200 unlabeled instances, and the 50th percentile is used to set the threshold.

3.1.3 Explaining Uncertainty

If the features used to explain uncertainty are identical to the features used to train the base model, the certainty function of x, x^*_C, can be used as the prediction probability function of certainty instead of introducing an intermediary model that predicts uncertainty[3]. In this case, the prediction probability functions of certainty are defined as follows:

$$P(Uncertain|x) = 1 - x^*_c$$
$$P(Certain|x) = x^*_c$$

When the features used for an explanation generation are different from the features used to train the base model, an intermediary model[4] can be used to predict the prediction certainty of the base model.

[1]For instance, a popular machine learning library *scikit-learn* provides APIs to estimate prediction probabilities for the most of the classification models in the library

[2]In fact, for this toy example, the features used for explanation are identical to the features used to train the base model. Therefore, explicit *uncertainty mapping* can be skipped and the probability prediction function of the base model can be directly used to formulate the prediction probability function needed for LIME.

[3]LIME only requires the prediction probability function of a model and does not require the actual model. Therefore, it is sufficient to provide the prediction probability function of certainty.

[4]The intermediary model must be fairly accurate in predicting certainties of the base model.

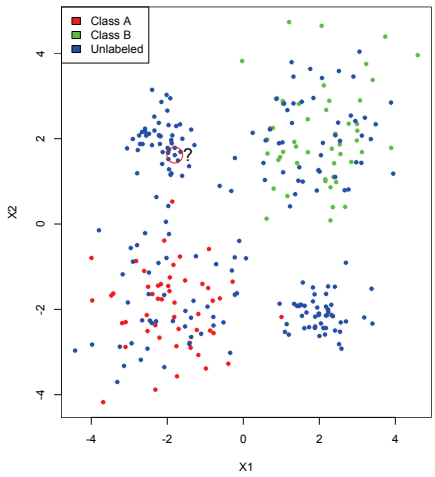

Figure 12: Query Instance

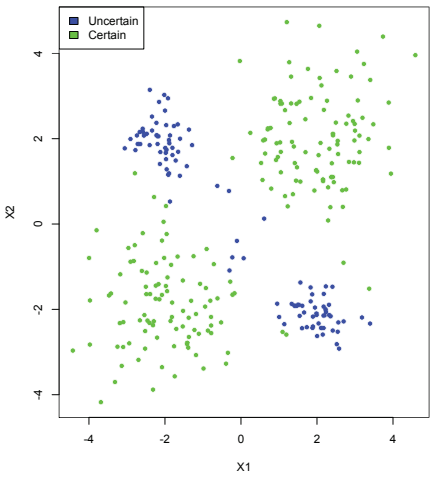

Figure 13: Uncertainty Mapping

Once *uncertainty mapping* is completed, an intermediary model is trained on the mapped data. Given an instance x, this model predicts whether the base model will accurately classify x regardless of the classification of x. Since this intermediary model predicts whether the prediction of an instance is certain or uncertain, when the prediction probability function of the intermediary model and a query instance are passed to LIME, LIME generates an explanation as to whether the query instance is either certain or uncertain. Since the query instance is most uncertain as long as the intermediary model is accurate, the prediction of the query instance will be uncertain.

For the toy example, an SVM with the radial basis function (RBF) kernel is used as the intermediary model and trained with all instances with the certainty labels obtained from *uncertainty mapping* (Figure 13). Then the intermediary model and q are given to LIME to generate an explanation of why q is uncertain.

The following is the explanation[5] provided by LIME for the query instance q with coordinate $(-1.80, 1.62)$ and Figure 14 shows the visualization of the explanation:

```
* X1 ≤ -0.16 (weight: uncertain)
* X2 > 0.59 (weight: uncertain)
```

Most of the certain instances are in Quadrant I $(+,+)$ and Quadrant III $(-,-)$ whereas most of the uncertain instances are in Quadrant II $(-,+)$ and Quadrant IV $(+,-)$. Thus, the explanation generated by LIME is an accurate description of why the active learner se-

[5]LIME also provides weight for each feature. For this example, both constraints contributed for the instance being uncertain.

lected q as the query instance. The active learner selected q because the model is uncertain when an instance is within the region defined by $X1 \leq -0.16$ and $X2 > 0.59$.

As discussed in Section 2.1.1, LIME provides a locally faithful explanation. As Figure 14 shows, the LIME's explanation does not apply to uncertain instances in Quadrant IV.

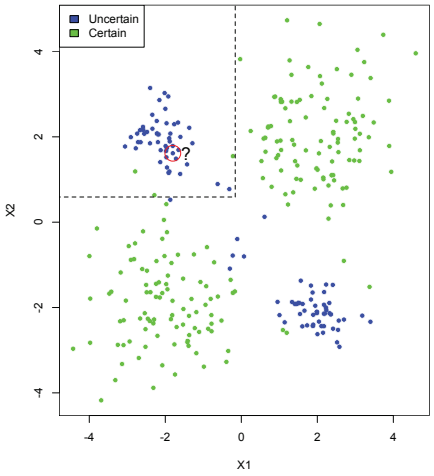

Figure 14: Explanation

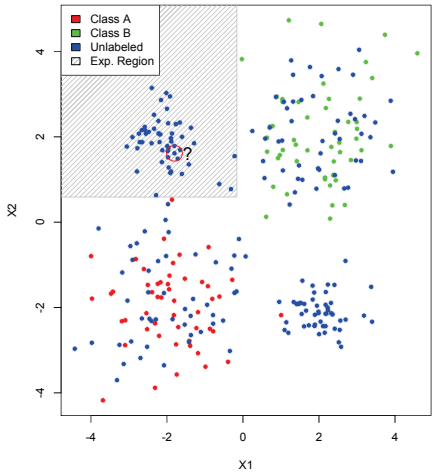

Figure 15: Explanatory Region

3.1.4 Explanatory Region

The LIME explanation of a prediction for tabular data consists of ranges of features (constraints). Thus, an explanation can be used to define an *explanatory region*, n dimension space where n is the number of constraints in the explanation.

For the toy example, LIME provides an explanation $X1 \leq -0.16$ and $X2 > 0.59$ for the query instance q. These two constraints define a region and Figure 15 illustrates the *explanatory region*.

3.2 Batch Selection

As discussed in Section 1.2.4, querying one instance at a time is often highly inefficient in practice. In a typical batch selection strategy, a set of Q instances is selected from the pool of unlabeled instances.

3.2.1 Explanatory Region-Based Batch Selection

Each active learning round can be considered as reducing or exploring the uncertain region (Section 1.1.1). The problem is that it is difficult for a human to understand what regions are being explored especially when the dataset has a large number of dimensions.

Given an explanatory region defined with constraints that contribute to an instance being uncertain, a different approach can be taken: selecting a batch from the explanatory region rather than the entire pool. By selecting a batch of unlabeled instances from an explanatory region, each active learning round explicitly explores the explanatory region.

The batch selected from an explanatory region is likely to have less variance in the batch compared to the other batch selected from the entire pool. Although this reduction in variance may lead to a performance loss, there are benefits to this batch selection strategy. First, a human annotator who labels the queries can have a better understanding of what labeling the batch means. Second, it can be easier to label a batch of instances that are similar to each other. For example, in the context of applying active learning to materials discovery, a chemist can have a better understanding of what performing a batch of reactions means since there is an explicit *explanatory region* defined. Moreover, it is easier to perform a batch of similar reactions compared to performing a batch of reactions that are very different.

The number of unlabeled instances that are within an explanatory region varies in size for different queries. It is possible that the number of unlabeled instances is smaller than the desired batch size. In order to cope with this, a bigger batch consisting of several smaller batches can be used.

For example, suppose an ideal batch size is 50. Then multiple batches selected from different explanatory regions can be used to fill the large batch. This process of creating a bigger batch out of smaller batches can be done in a greedy way. Algorithm 1 describes one possible approach.

3.2.2 K-Means Clustering Algorithm Based Batch Selection Strategy

As discussed in Section 1.2.4, *Q-best* is often not a good batch selection strategy and there have been studies on more effective batch selection strategies [6, 13]. Since the study of batch selection strategies is beyond the scope of the thesis and there is no standard implementation of those proposed batch selection strategies available, simple batch selection strategies based on *k-means clustering algorithm* are used.

The *k-means clustering algorithm* partitions n instances into k clusters, where the centroid of a cluster is defined as the mean of instances in the cluster [21]. Using the *k-means clustering algorithm* two strategies are made: *k-means uncertain* and *k-means closest*.

For both *k-means-uncertain* and *k-means-closest*, given a batch size k, partition the pool of unlabeled instances into k clusters. Then from each cluster, select a representative

Algorithm 1: Batch selection based on multiple explanatory regions

```
Let Q be the batch size
Let b be the maximum sub-batch size
Let Batch be an empty set
Let Exclusion be an empty set
while Batch size is less than Q do
    Get a query instance x
    Let Exp(x) be the explanation of uncertainty on x
    Let ExpRegion(x) be the region defined by Exp(x)
    Let ExpPool(x) be the unlabeled instances in ExpRegion(x) but not in
    Exclusion
    Select b instances from ExpPool(x) based on a batch selection strategy
    Add the selected instances to Batch
    Add the instances in ExpPool(x) to Exclusion
end
return Batch
```

instance: *k-means-uncertain* selects the one with lowest certainty value from the centroid, while *k-means-closest* selects the closest instance to the centroid. Algorithms 2 and 3 describe the details.

Algorithm 2: k-means-uncertain

```
Let P be the pool
Let Batch be an empty set
Let k be the batch size
Let C be the centroids from k-means clustering algorithm on P
Assign p ∈ P to a cluster represented by the closest centroid c ∈ C
for each cluster represented by c ∈ C do
    Let b be the instance that has the lowest certainty value in c
    Add b to Batch
end
return Batch
```

3.3 Applying the Technique to a Different Selection Strategy

As discussed in Section 1.3, there are several query selection strategies that are used in active learning. All the strategies mentioned in Section 1.3 select a query based on a certain measure whether it is *uncertainty sampling*, the QBC algorithm, or *expected error*

Algorithm 3: k-means-closest

```
Let P be the pool
Let Batch be an empty set
Let k be the batch size
Let C be the centroids from k-means clustering algorithm on P
for each cluster represented by c ∈ C do
    Let b be the closest instance to the centroid c.
    Add b to Batch
end
return Batch
```

reduction.

For instance, the QBC algorithm uses a disagreement coefficient to select a query. The QBC algorithm can be easily adapted to the technique by using disagreement coefficient instead of certainty. Instead of mapping certainty, we map disagreement. Once disagreement is mapped (i.e. is instances are labeled as either agreed or disagreed about the classification of the instances by the models in the committee) LIME can explain the disagreement of the query instance. Similarly, other selection strategies can be used.

4 Active Learning with Domain Knowledge Utilization

One of the problems with active learning is that the interaction between a human annotator and an active learner is one-directional. In a typical active learning setting, a human annotator's role is to provide the labels of the queries that the active learner asks for, and there is not much room to incorporate the annotator's domain knowledge.

In Section 3, I proposed a technique to explain active learning queries. An explanation of a query itself can be useful: it may provide insight to domain experts or it may allow domain experts to use their domain knowledge to detect a problem with the current learning strategy or model. Furthermore, I propose a hybrid active learning setting, in which experts incorporate their domain knowledge in a learning cycle. In this setting, the active learner recommends multiple queries along with explanations, and experts choose a query based on their interest and domain knowledge.

For instance, consider an active learning setting for materials discovery. In a typical active learning setting, chemists conduct reactions that the active learner wants. In this new setting, the active learner will recommend multiple reactions and provide explanations of why these reactions will improve the model, and chemists will choose reactions that they are more interested in, based on their domain knowledge. This query selection and labeling approach, in which domain knowledge is incorporated in the active learning cycle, may lead to a performance improvement.

An active learner does not have domain knowledge: although an active learner may think that knowing the label of an instance is likely to improve the model, the reason may be unsound or uninteresting to domain experts. Thus, explanations of queries can help experts make the decisions on whether to accept or reject labeling queries.

4.1 Incorporating Domain Knowledge in Active Learning Cycle

To demonstrate how domain knowledge can be incorporated in active learning and how it can increase learning rate, a synthetic dataset is generated for a simple binary classification problem.

4.1.1 Synthetic Data

The synthetic data consists of the four two-dimensional normal distributions with $\mu = (\pm 3, \pm 3)$ and $\sigma = 2$. For each distribution, 50 instances are randomly generated. Instances sampled from normal distributions centered at $\mu = (\pm 3, +3)$ are labeled as class A, while instances sampled from normal distributions centered at $\mu = (\pm 3, -3)$, are labeled as class B. Figure 16 shows all 200 instances.

For the training set, 20 instances are sampled: 10 from the normal distribution with $\mu = (+3, +3)$ and 10 from the normal distribution with $\mu = (-3, -3)$ (Figure 17). The training set is generated in a biased way: the distribution of training set is different from that of the population. The training set is generated in this way to mimic 'domain knowledge', which will be described in the next section.

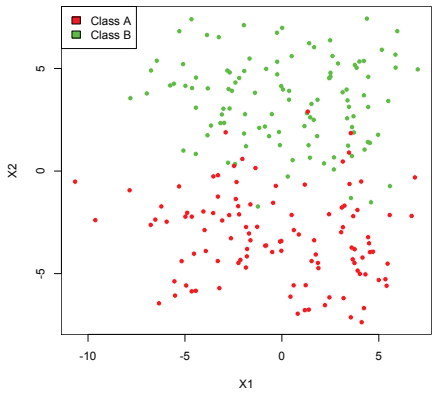

Figure 16: Synthetic Data

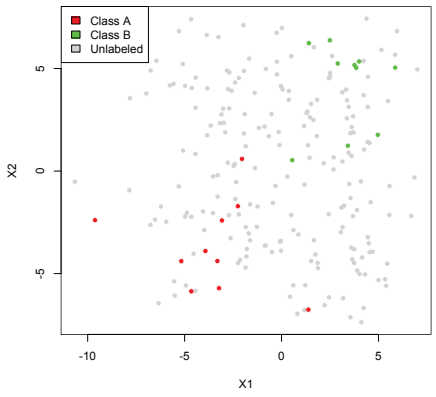

Figure 17: Training Set

4.1.2 Learning Cycle with Domain Knowledge

In the example, the training set is sampled from a subset of the population, which creates a sampling bias. We assume that an expert, although he or she does not know the population distribution, has full knowledge of the instances that are in the training set. In this setting, domain knowledge can be mimicked by making the artificial expert prefer queries that are substantially different from the instances that are already in the training set. Because of the bias in the training set, instances in Quadrant II and IV are more "interesting" to the domain expert.

For this toy example, in each round, the active learner recommends four potential queries based on *k-means-uncertain* (Algorithm 2). Among the four queries, the artificial expert selects a query to be labeled. To mimic domain knowledge, the artificial expert is encoded to prefer a query in Quadrant II or IV to a query in Quadrant I or III. If there is more than one query in Quadrant II or IV, one is randomly selected. If none of the four queries is in Quadrant II or IV, one of the four queries is randomly selected. By setting active learning in this way, both domain knowledge and explanations are simulated: where the explanations are what Quadrants instances are in.

For performance comparisons, active learning (uncertainty sampling) and passive learning (random sampling) are run with the same training set and pool.

4.1.3 Result

An SVM with degree 3 polynomial kernel is used as the model. The test set consists of 100 instances that are generated in the same way the pool is generated (25 from each of the four distributions). MCC is measured for each active learning round.

Figure 18 shows the average MCC over 100 trials. Overall, expert-aided active learning (EAL) outperforms both active learning (AL) and passive learning (PL). For the first 7 queries, the average MCCs of EAL are greater than those of AL at $\alpha = 0.01$ level (the t-test result can be found in Table 2 in Appendix A).

The toy example is devised in a meticulously controlled way. The purpose of the toy example is to show how domain knowledge can be incorporated in an active learning cycle and to show the potential performance benefit from it.

5 Expert-Aided Active Learning

In this section, expert-aided active learning will be tested on the Dark Reactions Project[6]. For each active learning round, a query explanation is provided to assist an expert in making a decision of accepting or rejecting a set of reactions recommended by the active learner.

[6]Code and data for replicating the experiments are available at `https://github.com/kh-chang/explaining-active-learning-queries`

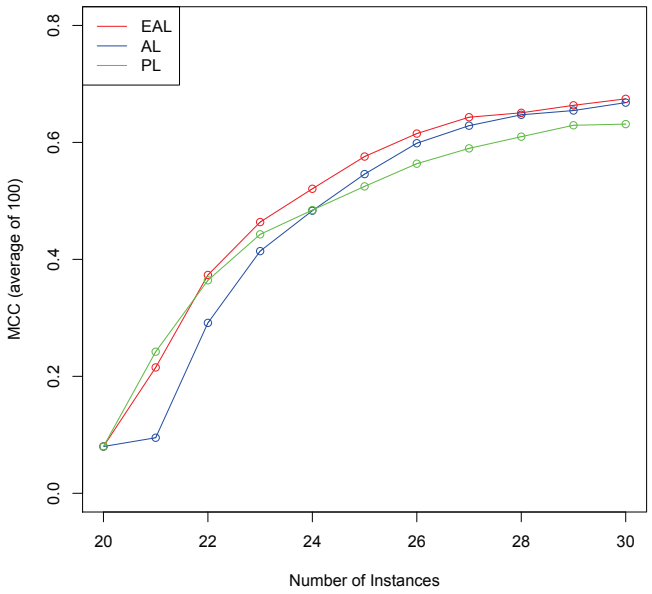

Figure 18: Incorporating Domain Knowledge in Active Learning

5.1 Dark Reactions Project Dataset

Haverford's the Dark Reactions Project (DRP) aims to help chemists to collect better data in a more efficient way by utilizing machine learning algorithms.

5.1.1 Description

The DRP dataset consists of organically-templated metal oxide synthesis reactions with information such as the reactants and their masses, temperature, pH, and time. The reaction is labeled as either a success or a failure in terms of the crystallinity of the reaction product. If a reaction produces a polycrystalline product, the reaction is considered as a success. The bar is intentionally set low because chemists have been good at fine-tuning a reaction that produces a polycrystalline product to the one that produces a single crystalline product. Such a fine-tuning process often involves trying numerous reactions with the same reactants but with varied masses, temperatures, etc. Moreover, conceptually the distance between polycrystalline and single crystalline is much shorter than the distance between no crystalline and single crystalline.

The entire space of reactions that the DRP covers is extremely large with dimension close to 270. The existing dataset, however, is nowhere close to covering the entire space.

Many reactions are similar in the process of fine-tuning mentioned above. Moreover, the reactions are chosen by chemists or the DRP recommendation system. As a result, three-quarters of the dataset are successful reactions. This is problematic: a classifier that classifies every reaction as a successful reaction will have an accuracy close to 75%. Thus, the DRP researchers suggested using more robust *Matthews correlation coefficient* (MCC), which takes into account true and false positives and negatives, for evaluating a machine learning model for the DRP dataset [22]. The MCC can be calculated from the confusion matrix using the formula [23]:

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

These characteristics of the DRP dataset make it difficult to work with it, but since the space is very large and mostly unexplored, and labeling a reaction means a chemist performing an actual experiment, the dataset is suitable to apply active learning.

5.1.2 Training and Test Sets

The specific version of the DRP dataset is from *Machine-learning-assisted materials discovery using failed experiments* [25]. The dataset consists of 3955 labeled reactions previously performed and additional 266 reactions. The 266 reactions are used as the validation set in Raccuglia et al.'s study [25]. The 266 reactions consist of 155 reactions recommended by the machine learning model in the paper and the rest 123 reactions are chosen by chemists. Since the goal of the DRP project is to predict the outcome of reactions with new combinations of reactants, the combinations of reactants in the 266 reactions are different from those of 3955 reactions.

To reflect this goal, the DRP researchers have been using a specially designed splitting method called *exploratory split* [22, 25] to split data into the training and test sets. While random split can potentially include the same combinations of reactants into both the training and test sets, *exploratory split* ensures that reactions with the same combination of reactants only appear in either the training set or the test set. This way of a split is more realistic since the model is used to predict the outcomes of reactions that have different combinations of reactants.

125 of 3955 reactions are missing values for the feature "slow cool". For this paper, these reactions are dropped, resulting in 3830 reactions. The 3830 reactions are then split into a training set and a test set using *exploratory split* (1/3-test and 2/3-training data split). The training set consists of 2585 (67%) instances and the test set consists of 1246 (33%) instances. The 266 reactions are used as the validation set.

5.2 Base Model for Active Learning

An SVM with the Pearson VII universal kernel (PUK) is one of the most effective machine learning models currently known for the DRP dataset [22]. The parameters used for PUK

are: $\sigma = 7, \omega = 0.5$ [25]. The same 273 features used in *Machine-learning-assisted materials discovery using failed experiments* are also used to train the base model [25].

5.3 Active Learning Scenario

5.3.1 Pool-based Active Learning Cycle

All labels of reactions in the training set are initially hidden from the active learner. The active learner queries instances from the pool of reactions whose outcomes are already known. In this way, a model validation can be done without the burden of actually performing reactions. On the other hand, this approach restricts the instances that the active learner can query. In application, a larger pool of unlabeled instances is desirable.

5.4 Does Active Learning Help?

5.4.1 Query Selection Strategy

For the SVM model with PUK, uncertainty sampling (Section 1.3.1) showed a promising result (Figure 19). To check the validity of uncertainty sampling, I trained the SVM model with the training set. Then each instance in the test set was assigned to a group based on its certainty value from the trained model. For each group, I computed the MCC of the model for the instances within the range of certainty. Figure 19 shows that the certainty of a prediction is in line with MCC computed based on the test set: the model indeed predicts instances with higher certainties better than the instances with lower certainties.

Since the model's prediction certainties are consistent with the actual performance of the model, uncertainty sampling (Section 1.3.1) is expected to be a suitable query selection strategy. The detailed result can be found in Appendix 3.

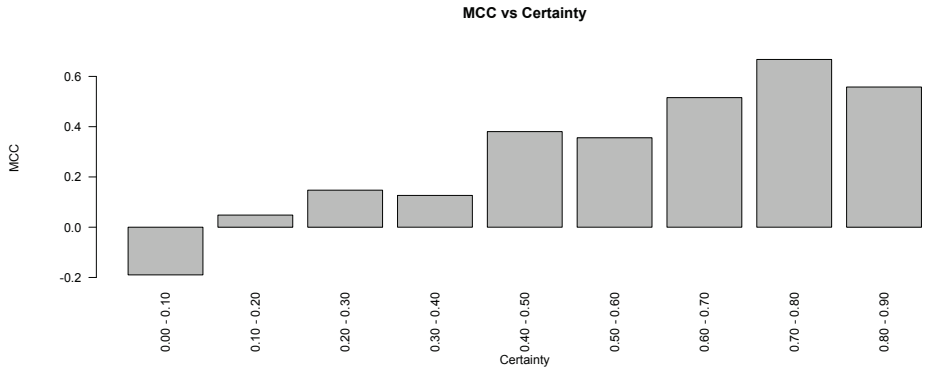

Figure 19: MCC versus Certainty

MCC tends to increase as certainty increases, suggesting that *uncertainty sampling* is a suitable query selection strategy for the DRP dataset with the SVM model.

5.4.2 Active Learning on the DRP Dataset

To validate that active learning is effective with the DRP dataset, a sequential (non-batch) active learning is tested against passive learning (random sampling). Figure 20 shows the average of MCCs of 10 trials. As the graph shows, active learning with uncertainty sampling clearly outperforms passive learning. There is a considerable gap between the MCCs of active learning and passive learning and the MCCs converge as the number of queries increases.

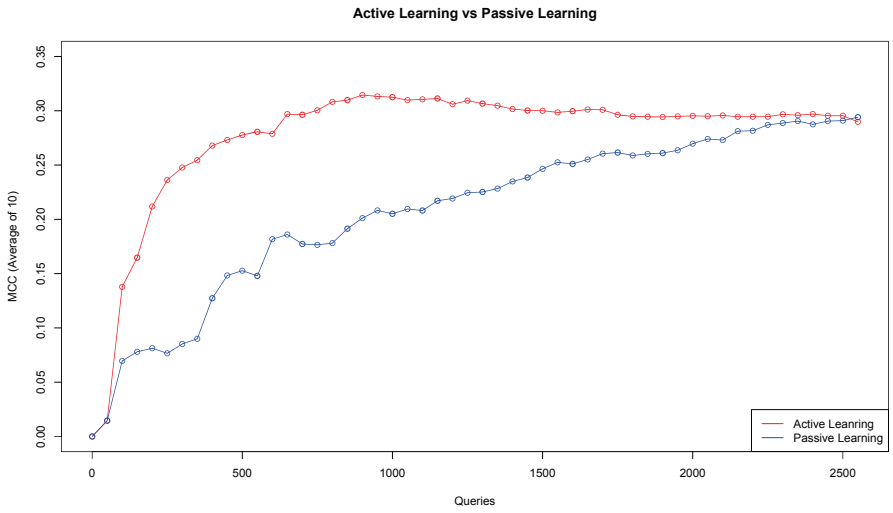

Figure 20: Active Learning on the DRP dataset

5.4.3 Batch Selection Strategy

Four different batch selection strategies are tested: *random sampling*, naive *Q-best* (Section 1.2.4), *k-means-uncertain* and *k-means-closest* (Section 3.2.2).

Figure 21 shows the MCCs of the four batch strategies on the DRP dataset with the batch size of 50. As the graph shows, *k-means-uncertain* outperforms other batch selection strategies. With *k-means-uncertain*, the active learner reached the MCC plateau with a considerably fewer number of instances. Therefore, *k-means-uncertain* is chosen as the batch strategy for the DRP dataset.

5.5 Model for Uncertainty Prediction

If the same set of features are used both for the base model and for explanation generation, there is no intermediary model needed. As discussed in Section 5.2, 273 features are used for the base model, but 62 features such as element names and element groups are excluded

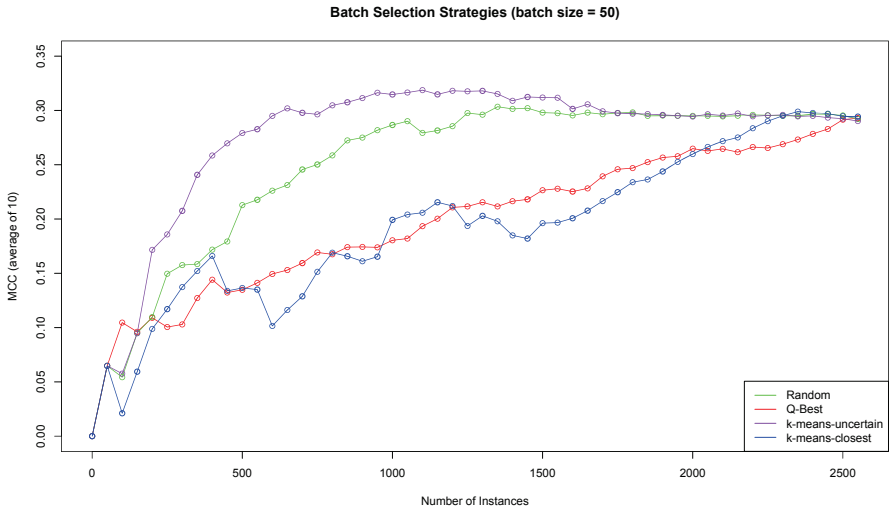

Figure 21: Batch Mode Active Learning on the DRP dataset

for explanation generation because they do not provide much additional information when a list of organic and inorganic reactants of reactions provided along with an explanation (Section 5.7.1). Since only a subset of features is used, an intermediary model is needed.

After *uncertainty mapping*, the intermediary model is trained on the mapped data to classify whether the base model will be certain or uncertain given an instance (Section 3.1.3). This model can be different from the base model that predicts whether a reaction will be successful. Four models are tested: SVM with PUK, Random Forest, *unweighted k-nearest neighbor*, and *weighted k-nearest neighbor*. Figure 22 shows model performances and the 30th percentile is used for *uncertainty mapping*. Overall, SVM with PUK shows a better performance than other models. Thus, SVM with PUK is chosen as the intermediary model that predicts uncertainty of the base model.

For all models tested, MCC decreases as the number of instances increases. Since the percentile for the uncertainty threshold is fixed at the 30th percentile, the number of uncertain instances decreases while the number of certain instances increases after *uncertainty mapping*, resulting in a very skewed distribution of certain and uncertain instances. Since the intermediary model becomes increasingly less accurate as the number of instances increases, a query explanation will also be less accurate as the number of instances increases. For the experiment, I concluded that it would not be appropriate to use the intermediary model beyond 500 instances since the intermediary model cannot accurately predict the certainty of predictions of the base model. In a typical active learning setting, the pool size is expected to be much greater than the training set size. Therefore, this issue of an intermediary model becoming increasingly less accurate is expected to be mitigated with

a greater pool size.

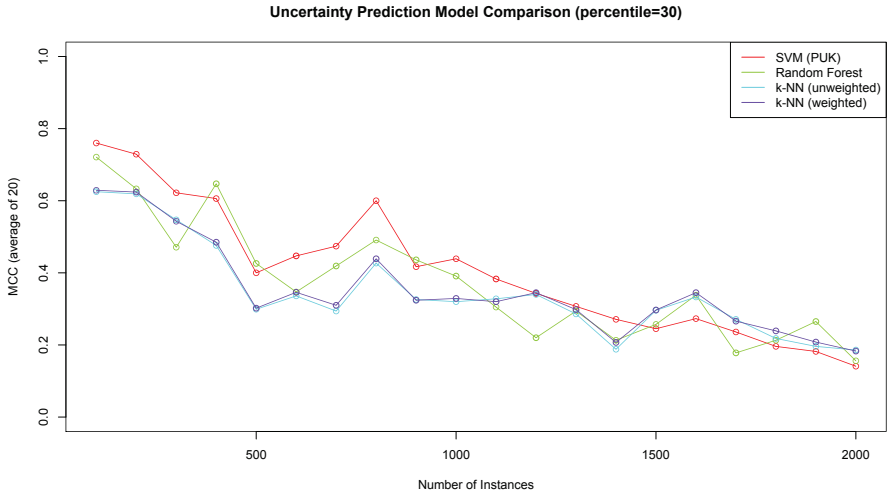

Figure 22: Uncertainty Prediction Model Comparison

5.6 Tuning the Parameters

There are several parameters that need to be determined: the batch size, the percentile for the threshold for *uncertainty mapping* (Section 3.1.2), and the number of constraints/features to explain an active learning query.

For the DRP dataset, the batch size is the number of chemical reactions to be performed for each active learning round. The batch size was set at 20 after a discussion with the DRP researchers. The percentile for the uncertainty threshold can be flexible as long as there is an enough mix of certain and uncertain instances after *uncertainty mapping*. The 30th percentile was used, which means that 30 percent of the pool (unlabeled instances) will be mapped as uncertain.

A LIME explanation consists of a set of constraints that defines an *explanatory region*. Having too few constraints leads to uninteresting explanations (e.g. a single constraint, `pH < 5`, is uninteresting and results in an *explanatory region* that is too large). The ideal number of constraints is the one that provides enough information as to why the model is uncertain about the query instance but does not create a much cognitive burden to an expert. After a discussion with the DRP researchers, the number of constraints explaining uncertainty is set to range from 4 to 6.

5.7 Query Explanation

Along with the uncertainty explanation of a query instance, a description of the query instance as well as a description of the batch is provided to the expert.

5.7.1 Query Instance Description

```
* inorg1       : Oxovanadium(2+) sulfate
* inorg1mass   : 0.1648 grams
* inorg2       : Selenium dioxide
* inorg2mass   : 1.2995 grams
* org1         : 1,4-dimethylpiperazine
* org1mass     : 0.1166 grams
* temp         : 110
* time         : 24.0
* slowCool     : True
* pH           : 4
```

Figure 23: Query Instance Description

It is important to provide succinct information about the query instance to chemists. In this way, a chemist can have a concrete understanding of what the query instance is. An instance description of a query reaction lists the reactants and their masses in the reaction along with other information such as temperature, time, a logical value for slow cool and pH. This information is what a chemist enters into the DRP database and the rest of features used in the training set are generated from this information. Figure 23 shows a query instance description.

5.7.2 Uncertainty Explanation

```
* 1.76 < PaulingElectronegGeom <= 1.89
* 1.76 < PaulingElectronegMean <= 1.91
* orgminimalprojectionradiusMax <= 2.84
* slowCool = True
* orghbdamsaccGeomAvg <= 0.99
* numberInorg = 2
```

Figure 24: Uncertainty Explanation
LIME generated feature constraints that contribute to the query instance being uncertain.

A LIME explanation provides constraints of features with weights whether a feature contributes to classifying a given instance as one class or another. Since the purpose of

using LIME is to explain the uncertainty of a query instance for the DRP dataset, given a set of constraints, only the constraints contributing to the query instance being uncertain are used. In short, an explanation of a query reaction is an answer to a chemist's question "why should I perform this reaction?" Figure 24 shows the uncertainty explanation of the query instance in Figure 23.

5.7.3 Batch Description

```
* inorg1       : Oxovanadium(2+) sulfate,
                 sodium tellurite,
                 tellurium dioxide,
                 sodium metavanadate
* inorg1mass   : [0.0292, 0.4345] grams
* inorg2       : sodium tellurite,
                 ammonium metavanadate,
                 tellurium dioxide,
                 sodium metavanadate
* inorg2mass   : [0.0294, 0.4062] grams
* org1         : N,N'-dimethylethylenediamine,
                 ethylenediamine
* org1mass     : [0.0129, 0.2346] grams
* temp         : [90.0, 110.0]
* time         : [17.0, 60.0]
* pH           : [1.0, 8.0]
```

Figure 25: Batch Description
The batch description shows the variation in the batch of reactions. For example, in this batch, organic reaction is either N,N'-dimethylethylenediamine or ethylenediamine and its mass varies from 0.0128 to 0.2346 grams

An uncertainty explanation defines an explanatory region. As discussed in Section 3.2, a batch of reactions is selected based on the explanatory region of a query reaction. Since the dimension is very high for the DRP dataset, it is important to provide a description of the batch of reactions. This description aims at helping chemists understand variation in the batch of reactions. Figure 25 shows the description of 20 instances that are selected from the explanatory region defined by the constraints in Figure 24 using the *k-means-uncertain* batch selection strategy.

5.8 Experiment Setup

The experiment is set up to evaluate the understandability of query explanations and test whether incorporating domain knowledge in the active learning cycle leads to a performance improvement. The software used for the experiment is implemented using *scikit-learn* and *libact* [7, 37]. 100 reactions are randomly selected from the training set as a starting point for active learning. As discussed earlier, the batch size is set at 20 and the experiment consists of 20 active learning rounds. An explanatory-region-based batch selection algorithm (Algorithm 1) that uses *k-means-uncertain* as a sub-batch selection strategy is used[7]. Alex Norquist, Associate Professor of Chemistry at Haverford College, served as the expert.

The active learner recommends a set of reactions with explanation described in Section 5.7. Based on domain knowledge and explanation provided, the expert either accepts or rejects the set of reaction. The experiment took about an hour. During the 20 active learning rounds, a total of 43 questions are asked to the expert focusing on whether to accept and perform reactions or to reject the reactions. Among 43 questions, the expert accepted 31 and rejected 12. The raw output of the experiment is available in Appendix D.

5.9 Result

Ideally, the experiment is repeated multiple times to compare the performance of EAL with those of active learning (AL) and passive learning (PL). However, since a single run of EAL takes an hour of human time and requires an expert in chemistry, the experiment was run once. For each round, the model was tested against the test set and the validation set. Since AL and PL can be easily repeated, AL and PL are run 20 times. Figures 26 and 27 show the MCCs for each round. For both the test set and the validation set, EAL shows a commensurate MCC with the average MCC of AL and a greater MCC than the average MCC of PL. Since the EAL was only run once, there is not much of statistical significance of the result in terms of performance gain. The numeric values of Figures 26 and 27 can be found in Appendix C.

Regarding providing interpretable explanations of active learning queries, explanations were interpretable enough to Norquist who served as the chemistry expert for this experiment. He was able to make some use of the explanations based on his domain knowledge to decide whether to accept or reject a set of recommended reactions. He commented that the explanations were helpful to identify the parts of chemical space that are poorly understood by the model. He also pointed out the cases where an explanation and a batch of reactions do not match. For example, an explanation only uses organic features but the batch of reactions varies in inorganic reactants.

[7] Both Q and b are set at 20.

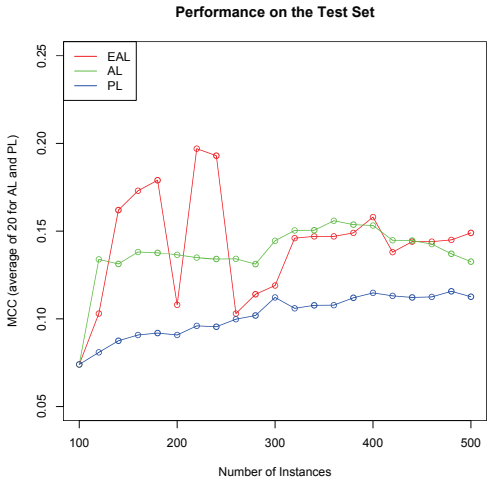

Figure 26: Performance on the Test Set

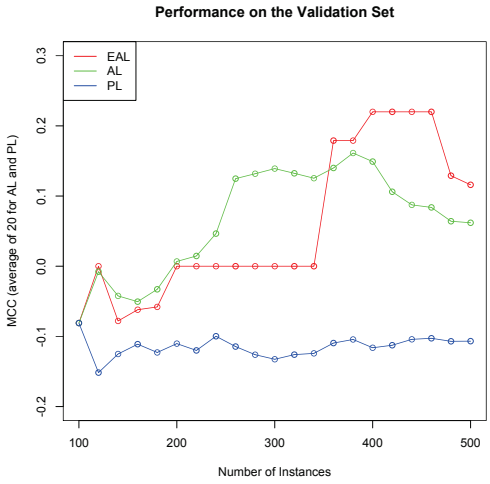

Figure 27: Performance on the Validation Set

6 Conclusion and Future Work

In this thesis, I was able to produce interpretable explanations for active learning queries based on certainties derived from the prediction probability function of the base model and using LIME to explain uncertainties of query instances,. For the DRP dataset, the expert was able to interpret the explanation and use it to understand how the model behaves. Incorporating domain knowledge in active learning showed a potential: the performance of expert-aided active learning (EAL) was comparable to the average performance of AL on both the training and the test set. However, because EAL requires a domain expert and considerable human time the experiment could not be repeated multiple times to draw a statistical significance out of the result. A more through performance comparison is left for future work.

As discussed in Section 3.1.3, when the set of features used to train the model and the set of features used for explanation generation differ, an intermediary model was used. This approach, however, introduces a layer of potential error, and it is certainly problematic if the intermediary model cannot accurately predict the uncertainties of the base model. An approach that does not require the intermediary model is desirable.

As the chemistry expert, Norquist, pointed out sometimes there was inconsistency between the uncertainty explanation and the variation existing in the batch. This issue arose from using locally faithful constraints to define an explanatory region. Therefore, although all reactions in the batch are uncertain reactions, not all the reactions are uncertain for the same reason. A more sophisticated batch selection strategy based on an *explanatory region* that results in a set of reactions which are consistent with uncertainty explanation

is another possible direction for a future study.

7 Acknowledgment

This thesis could not have been completed without the guidance and full support from my advisor Sorelle Friedler. I was very lucky to have a devoted advisor like Sorelle. Richard Phillips provided enormous support. He was always willing to discuss the ideas with me and gave thoughtful feedback.

I would like to thank Alex Norquist, Associate Professor of Chemistry, and Joshua Shrier, Associate Professor of Chemistry. Alex Norquist very generously spent his time testing expert-aided active learning. Joshua Schrier's feedback was critical in formatting the explanations in the way described in Section 5.7.

I would also like to thank Tian Jin and Jeff Son for helpful discussions and feedback, and Sang Dae Suh for proofreading.

References

[1] Dana Angluin. Queries and concept learning. *Machine learning*, 2(4):319–342, 1988.

[2] Les E. Atlas, David A. Cohn, and Richard E. Ladner. Training connectionist networks with queries and selective sampling. In D. S. Touretzky, editor, *Advances in Neural Information Processing Systems 2*, pages 566–573. Morgan-Kaufmann, 1990.

[3] Jason Baldridge and Miles Osborne. Active learning and the total cost of annotation. In *EMNLP*, pages 9–16, 2004.

[4] Nahla Barakat and Joachim Diederich. Eclectic rule-extraction from support vector machines. *International Journal of Computational Intelligence*, 2(1):59–62, 2005.

[5] Eric B Baum and Kenneth Lang. Query learning can work poorly when a human oracle is used. In *International Joint Conference on Neural Networks*, volume 8, page 8, 1992.

[6] Klaus Brinker. Incorporating diversity in active learning with support vector machines. In *ICML*, volume 3, pages 59–66, 2003.

[7] Lars Buitinck, Gilles Louppe, Mathieu Blondel, Fabian Pedregosa, Andreas Mueller, Olivier Grisel, Vlad Niculae, Peter Prettenhofer, Alexandre Gramfort, Jaques Grobler, Robert Layton, Jake VanderPlas, Arnaud Joly, Brian Holt, and Gaël Varoquaux. API design for machine learning software: experiences from the scikit-learn project. In *ECML PKDD Workshop: Languages for Data Mining and Machine Learning*, pages 108–122, 2013.

[8] David A Cohn, Zoubin Ghahramani, and Michael I Jordan. Active learning with statistical models. *Journal of artificial intelligence research*, 1996.

[9] Mark W Craven and Jude W Shavlik. Extracting tree-structured representations of trained networks. *Advances in neural information processing systems*, pages 24–30, 1996.

[10] Sanjoy Dasgupta. Two faces of active learning. *Theor. Comput. Sci.*, 412(19):1767–1781, April 2011.

[11] Sanjoy Dasgupta and Daniel Hsu. Hierarchical sampling for active learning. In *Proceedings of the 25th International Conference on Machine Learning*, ICML '08, pages 208–215, New York, NY, USA, 2008. ACM.

[12] Thomas G. Dietterich. Ensemble methods in machine learning. In *Proceedings of the First International Workshop on Multiple Classifier Systems*, MCS '00, pages 1–15, London, UK, 2000. Springer-Verlag.

[13] Yuhong Guo and Dale Schuurmans. Discriminative batch mode active learning. In *Advances in neural information processing systems*, pages 593–600, 2008.

[14] Steve Hanneke. Theory of active learning. Computer sciences technical report, 2014. `http://www.stevehanneke.com`.

[15] Rebecca Hwa. On minimizing training corpus for parser acquisition. In *Proceedings of the 2001 Workshop on Computational Natural Language Learning - Volume 7*, ConLL '01, pages 10:1–10:6, Stroudsburg, PA, USA, 2001. Association for Computational Linguistics.

[16] Ajay J. Joshi, Fatih Porikli, and Nikolaos Papanikolopoulos. Multi-class active learning for image classification. In *2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2009), 20-25 June 2009, Miami, Florida, USA*, pages 2372–2379, 2009.

[17] Ross D King, Kenneth E Whelan, Ffion M Jones, Philip GK Reiser, Christopher H Bryant, Stephen H Muggleton, Douglas B Kell, and Stephen G Oliver. Functional genomic hypothesis generation and experimentation by a robot scientist. *Nature*, 427(6971):247–252, 2004.

[18] David D. Lewis and Jason Catlett. Heterogeneous uncertainty sampling for supervised learning. In *In Proceedings of the Eleventh International Conference on Machine Learning*, pages 148–156. Morgan Kaufmann, 1994.

[19] David D. Lewis and William A. Gale. A sequential algorithm for training text classifiers. In *Proceedings of the 17th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '94, pages 3–12, New York, NY, USA, 1994. Springer-Verlag New York, Inc.

[20] Charles X. Ling and Jun Du. Active learning with direct query construction. In *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '08, pages 480–487, New York, NY, USA, 2008. ACM.

[21] Stuart Lloyd. Least squares quantization in pcm. *IEEE transactions on information theory*, 28(2):129–137, 1982.

[22] Geoffrey Martin-Noble. Optimizing a machine learning system for materials discovery. Bachelor thesis, Haverford, 2016.

[23] Brian W Matthews. Comparison of the predicted and observed secondary structure of t4 phage lysozyme. *Biochimica et Biophysica Acta (BBA)-Protein Structure*, 405(2):442–451, 1975.

[24] Tom Michael Mitchell. Version spaces: an approach to concept learning. Technical report, DTIC Document, 1978.

[25] Paul Raccuglia, Katherine C. Elbert, Philip D. F. Adler, Casey Falk, Malia B. Wenny, Aurelio Mollo, Matthias Zeller, Sorelle A. Friedler, Joshua Schrier, and Alexander J. Norquist. Machine-learning-assisted materials discovery using failed experiments. *Nature*, 533(7601):73–76, 05 2016.

[26] Marco Túlio Ribeiro. Lime. `https://raw.githubusercontent.com/marcotcr/lime/master/doc/images/lime.png`, 2016.

[27] Marco Túlio Ribeiro, Sameer Singh, and Carlos Guestrin. "Why should I trust you?": Explaining the predictions of any classifier. *CoRR*, abs/1602.04938, 2016.

[28] Tobias Scheffer, Christian Decomain, and Stefan Wrobel. Active hidden markov models for information extraction. In *International Symposium on Intelligent Data Analysis*, pages 309–318. Springer, 2001.

[29] Burr Settles. Active learning literature survey. Computer Sciences Technical Report 1648, University of Wisconsin–Madison, 2009.

[30] Burr Settles. *Active Learning*. Synthesis lectures on artificial intelligence and machine learning. Morgan & Claypool, 2011.

[31] Burr Settles, Mark Craven, and Soumya Ray. Multiple-instance active learning. In J. C. Platt, D. Koller, Y. Singer, and S. T. Roweis, editors, *Advances in Neural Information Processing Systems 20*, pages 1289–1296. Curran Associates, Inc., 2008.

[32] H. S. Seung, M. Opper, and H. Sompolinsky. Query by committee. In *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*, COLT '92, pages 287–294, New York, NY, USA, 1992. ACM.

[33] Claude Elwood Shannon. A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communications Review*, 5(1):3–55, 2001.

[34] Katrin Tomanek and Udo Hahn. Semi-supervised active learning for sequence labeling. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP: Volume 2 - Volume 2*, ACL '09, pages 1039–1047, Stroudsburg, PA, USA, 2009. Association for Computational Linguistics.

[35] Manfred K. Warmuth, Jun Liao, Gunnar Rtsch, Michael Mathieson, Santosh Putta, and Christian Lemmen. Support vector machines for active learning in the drug discovery process. *Journal of Chemical Information Sciences*, 43:667–673, 2003.

[36] Manfred K Warmuth, Gunnar Rätsch, Michael Mathieson, Jun Liao, and Christian Lemmen. Active learning in the drug discovery process. In T. G. Dietterich, S. Becker, and Z. Ghahramani, editors, *Advances in Neural Information Processing Systems 14*, pages 1449–1456. MIT Press, 2002.

[37] Yao-Yuan Yang, Shao-Chuan Lee, Yu-An Chung, Tung-En Wu, Si-An Chen, and Hsuan-Tien Lin. libact: Pool-based active learning in python. Technical report, 2015.

[38] Zhipeng Ye, Peng Liu, Jiafeng Liu, Xianglong Tang, and Wei Zhao. Practice makes perfect. *Neurocomput.*, 196(C):95–106, July 2016.

[39] Xiaojin Zhu, John Lafferty, and Ronald Rosenfeld. *Semi-supervised learning with graphs*. Carnegie Mellon University, language technologies institute, school of computer science, 2005.

Appendix A Incorporating Domain Knowledge in Active Learning

Table 1: Average MCCs of 20 Runs

	21	22	23	24	25	26	27	28	29	30
EAL	0.215	0.373	0.464	0.521	0.576	0.615	0.643	0.651	0.663	0.674
AL	0.095	0.291	0.414	0.483	0.546	0.599	0.629	0.647	0.654	0.668
PL	0.242	0.364	0.443	0.484	0.525	0.564	0.590	0.610	0.629	0.631

Hypothesis: true means of MCCs of EAL and AL are equal
Alternative Hypothesis: true mean of MCCs of EAL is greater than that of AL
Degrees of Freedom: 99

Table 2: p-values for the t-tests

Number of Instances	n-th Query	p-value
21	1	8.66121e-09
22	2	1.64365e-07
23	3	3.82987e-05
24	4	3.90330e-05
25	5	1.94996e-06
26	6	0.0041851
27	7	0.0007130
28	8	0.1149893
29	9	0.0018019
30	10	0.0150918

Appendix B Query Selection Strategy on the DRP Dataset

Table 3: MCC vs Certainty

Certainty	TN	FN	TP	FP	Total	MCC
0.00 - 0.10	15	8	3	14	40	-0.19
0.10 - 0.20	9	5	19	27	60	0.05
0.20 - 0.30	9	4	16	18	47	0.15
0.30 - 0.40	8	5	19	17	49	0.13
0.40 - 0.50	14	7	23	9	53	0.38
0.50 - 0.60	20	11	48	18	97	0.36
0.60 - 0.70	30	8	84	24	146	0.52
0.70 - 0.80	96	15	564	62	737	0.67
0.80 - 0.90	1	0	28	2	31	0.56

Appendix C Active Learning on the DRP Dataset

Table 4: Performance on the Test Set

	120	140	160	180	200	220	240	260	280	300
EAL	0.103	0.162	0.173	0.179	0.108	0.197	0.193	0.103	0.114	0.119
AL	0.134	0.131	0.138	0.138	0.137	0.135	0.134	0.134	0.131	0.144
PL	0.081	0.087	0.091	0.092	0.091	0.096	0.096	0.100	0.102	0.112
	320	340	360	380	400	420	440	460	480	500
EAL	0.146	0.147	0.147	0.149	0.158	0.138	0.144	0.144	0.145	0.149
AL	0.150	0.150	0.156	0.154	0.153	0.145	0.145	0.143	0.137	0.133
PL	0.106	0.108	0.108	0.112	0.115	0.113	0.112	0.113	0.116	0.113

For AL and PL, MCCs are the average of 20 runs.

Table 5: Performance on the Validation Set

	120	140	160	180	200	220	240	260	280	300
EAL	0.000	-0.078	-0.062	-0.058	0.000	0.000	0.000	0.000	0.000	0.000
AL	-0.008	-0.042	-0.051	-0.033	0.007	0.015	0.046	0.125	0.132	0.139
PL	-0.152	-0.125	-0.111	-0.123	-0.110	-0.120	-0.100	-0.115	-0.126	-0.133
	320	340	360	380	400	420	440	460	480	500
EAL	0.000	0.000	0.179	0.179	0.220	0.220	0.220	0.220	0.129	0.116
AL	0.132	0.125	0.140	0.161	0.149	0.106	0.087	0.084	0.064	0.062
PL	-0.126	-0.124	-0.110	-0.104	-0.116	-0.113	-0.104	-0.103	-0.107	-0.107

For AL and PL, MCCs are the average of 20 runs.

Appendix D Experiment Output

```
================================================================================
Round 1
================================================================================
Explaining Query with id #62
Certainty 0.000

========== The model is uncertain about the reaction because ... ==============
 * slowCool = True
 * orgASA_PMin <= 13.92
 * 3.24 < orgpolarsurfaceareaMin <= 8.28
 * 13.99 < orgavgpolMax <= 14.53

Instances in the batch: 20
================ A potential variation for probing the region ==================
 * inorg1       : Oxovanadium(2+) sulfate, sodium vanadium trioxide,
                  ammonium metavanadate
 * inorg1mass   : [0.12, 1.899] grams
 * inorg2       : Selenium dioxide
 * inorg2mass   : [0.4397, 1.3726] grams
 * org1         : 1,4-dimethylpiperazine
 * org1mass     : [0.11, 0.8351] grams
 * temp         : [110.0, 120.0]
 * time         : [18.0, 24.0]
 * pH           : [4.0, 7.0]

============================== An example reaction ==============================
 * inorg1       : Oxovanadium(2+) sulfate
 * inorg1mass   : 0.1648 grams
 * inorg2       : Selenium dioxide
 * inorg2mass   : 1.2995 grams
 * org1         : 1,4-dimethylpiperazine
 * org1mass     : 0.1166 grams
 * temp         : 110
 * time         : 24.0
 * slowCool     : 1
 * pH           : 4

Would you like to conduct the experiments for this region? (y/n) : y
```

```
================================================================================
Round 2
================================================================================
Explaining Query with id #91
Certainty 0.000

=========== The model is uncertain about the reaction because ... ==============
 * inorg-org-moleratio <= -1.81
 * orgdon-wateraccratio > 0.04
 * slowCool = True
 * orgASA-Max > 102.13
 * numberInorg = 2
 * numberComponents >= 4

Instances in the batch: 3
================ A potential variation for probing the region ==================
 * inorg1       : Oxovanadium(2+) sulfate
 * inorg1mass   : [0.1686, 0.1736] grams
 * inorg2       : Selenium dioxide
 * inorg2mass   : [1.3456, 1.419] grams
 * org1         : 1,10-diaminodecane
 * org2         : 2,5-dimethylpiperazine
 * org2mass     : [0.1166, 0.1237] grams
 * pH           : [1.0, 7.0]

============================== An example reaction ==============================
 * inorg1       : Oxovanadium(2+) sulfate
 * inorg1mass   : 0.1686 grams
 * inorg2       : Selenium dioxide
 * inorg2mass   : 1.419 grams
 * org1         : 1,10-diaminodecane
 * org2         : 2,5-dimethylpiperazine
 * org2mass     : 0.1166 grams
 * temp         : 110
 * time         : 24.0
 * slowCool     : 1
 * pH           : 1

Would you like to conduct the experiments for this region? (y/n) : n

Explaining Query with id #170
```

```
Certainty 0.000

========== The model is uncertain about the reaction because ... ==============
 * slowCool = True
 * numberInorg = 2
 * 80.17 < orgASA-Min <= 98.46
 * 77.12 < orgASA-ArithAvg <= 98.46
 * orgASA_HArithAvg > 267.18

Instances in the batch: 20
================ A potential variation for probing the region ==================
 * inorg1       : Oxovanadium(2+) sulfate, sodium vanadium trioxide,
                  ammonium metavanadate
                  sodium metavanadate
 * inorg1mass   : [0.0233, 0.3357] grams
 * inorg2       : sodium tellurite, Selenium dioxide
 * inorg2mass   : [0.0815, 1.765] grams
 * org1         : 2,6-dimethylpiperazine, N,N-dimethyl-N'-ethylethylenediamine
                  3-aminoquinuclidine dihydrochloride
                  1,8-diaminooctane, 2,5-dimethylpyrazine
 * org1mass     : [0.0214, 0.3876] grams
 * temp         : [90.0, 150.0]
 * time         : [20.0, 60.0]
 * pH           : [1.0, 8.0]

============================== An example reaction ==============================
 * inorg1       : ammonium metavanadate
 * inorg1mass   : 0.0457 grams
 * inorg2       : Selenium dioxide
 * inorg2mass   : 0.3242 grams
 * org1         : 2,5-dimethylpyrazine
 * org1mass     : 0.0265 grams
 * temp         : 90
 * time         : 20.0
 * slowCool     : 1
 * pH           : 7

Would you like to conduct the experiments for this region? (y/n) : y

================================================================================
Round 3
```

```
================================================================================
Explaining Query with id #45
Certainty 0.000

========== The model is uncertain about the reaction because ... ==============
 * orghbdamsaccMax > 5.49
 * PaulingElectronegMean <= 1.64
 * 44.42 < EAGeom <= 53.82
 * slowCool = True
 * orgASA_HMax > 247.72

Instances in the batch: 20
================ A potential variation for probing the region ==================
 * inorg1       : Oxovanadium(2+) sulfate, vanadium(V) oxide
 * inorg1mass   : [0.0781, 0.8472] grams
 * org1         : H3PO3
 * org1mass     : [0.2127, 1.0613] grams
 * org2         : N,N'-dimethylethylenediamine, 1,8-diaminooctane,
                  tetraethylenepentamine, 2-methylpiperazine,
                  triethylenetetramine, N,N,N',N'-Tetramethylethylenediamine
                  N,N-dimethyl-N'-ethylethylenediamine
 * org2mass     : [0.1318, 0.9367] grams
 * temp         : [110.0, 180.0]
 * time         : [24.0, 120.0]
 * pH           : [2.0, 8.0]

============================== An example reaction ==============================
 * inorg1       : vanadium(V) oxide
 * inorg1mass   : 0.2035 grams
 * org1         : H3PO3
 * org1mass     : 1.0525 grams
 * org2         : N,N,N',N'-Tetramethylethylenediamine
 * org2mass     : 0.2565 grams
 * temp         : 180
 * time         : 48.0
 * slowCool     : 1
 * pH           : 2

Would you like to conduct the experiments for this region? (y/n) : y

================================================================================
```

```
Round 4
================================================================================
Explaining Query with id #202
Certainty 0.000

========== The model is uncertain about the reaction because ... ==============
 * numberOrg >= 2
 * 79.96 < orgASA-Max <= 141.13
 * 1.76 < PaulingElectronegGeom <= 1.87
 * 237.45 < orgASA_HMax <= 335.55
 * 352.67 < PearsonElectronegMean <= 446.73
 * 6.14 < orgASA_PGeomAvg <= 43.17

Instances in the batch: 5
================ A potential variation for probing the region ==================
 * inorg1       : vanadium(V) oxide
 * inorg1mass   : [0.1808, 0.1844] grams
 * inorg2       : sodium tellurite
 * inorg2mass   : [0.1143, 0.1171] grams
 * org1         : 1,6-diaminohexane, 3-aminoquinuclidine dihydrochloride
 * org1mass     : [0.123, 0.3977] grams
 * org2         : ethanol
 * org2mass     : [2.0573, 4.0394] grams
 * time         : [120.0, 144.0]
 * pH           : [4.0, 11.0]

============================== An example reaction ==============================
 * inorg1       : vanadium(V) oxide
 * inorg1mass   : 0.1829 grams
 * inorg2       : sodium tellurite
 * inorg2mass   : 0.1171 grams
 * org1         : 3-aminoquinuclidine dihydrochloride
 * org1mass     : 0.2002 grams
 * org2         : ethanol
 * org2mass     : 3.9743 grams
 * temp         : 110
 * time         : 144.0
 * slowCool     : 0
 * pH           : 7

Would you like to conduct the experiments for this region? (y/n) : y
```

```
Explaining Query with id #390
Certainty 0.000

========== The model is uncertain about the reaction because ... ==============
 * PaulingElectronegMean <= 1.64
 * PaulingElectronegGeom <= 1.64
 * 34.33 < EAMean <= 54.03
 * 338.58 < PearsonElectronegGeom <= 352.66
 * 33.57 < EAGeom <= 53.82

Instances in the batch: 15
================ A potential variation for probing the region ==================
 * inorg1       : Oxovanadium(2+) sulfate, sodium vanadium trioxide,
                  ammonium metavanadate, vanadium(V) oxide,
                  potassium vanadium trioxide
 * inorg1mass   : [0.0488, 1.756] grams
 * inorg2       : ammonium fluoride
 * org1         : H3PO3, 2-methylpiperazine,
                  3-aminoquinuclidine dihydrochloride,
                  N,N'-dimethylethylenediamine, ethylenediamine,
                  1,4-dimethylpiperazine, imidazole
 * org1mass     : [0.1182, 1.0235] grams
 * org2         : diethylenetriamine, ethylenediamine, 1,3-diaminopropane
                  1,2-diaminopropane, N,N,N',N'-Tetramethylethylenediamine
 * org2mass     : [0.1105, 0.5995] grams
 * temp         : [90.0, 180.0]
 * time         : [24.0, 120.0]
 * pH           : [2.0, 8.0]

============================== An example reaction ==============================
 * inorg1       : potassium vanadium trioxide
 * inorg1mass   : 0.1564 grams
 * org1         : ethylenediamine
 * org1mass     : 0.5644 grams
 * temp         : 110
 * time         : 24.0
 * slowCool     : 1
 * pH           : 5

Would you like to conduct the experiments for this region? (y/n) : n
```

```
Explaining Query with id #771
Certainty 0.000

========== The model is uncertain about the reaction because ... ==============
 * slowCool = True
 * orgASA_HArithAvg > 247.55
 * 237.45 < orgASA_HMax <= 335.55
 * orgASA_HGeomAvg > 247.47

Instances in the batch: 15
================ A potential variation for probing the region ==================
 * inorg1       : Oxovanadium(2+) sulfate, gallium trinitrate,
                  sodium vanadium trioxide, ammonium metavanadate,
                  sodium metavanadate, sodium fluoride,
                  molybdenum trioxide
 * inorg1mass   : [0.0865, 1.638] grams
 * inorg2       : Potassium Dichromate, Selenium dioxide
 * inorg2mass   : [0.0628, 1.346] grams
 * inorg3       : Selenium dioxide
 * org1         : 2-methylpiperazine, 2,5-dimethylpiperazine, 1,8-diaminooctane,
                  1,4-bis(3-aminopropyl)piperazine,
                  N,N'-dimethylethylenediamine, 3-aminoquinuclidine,
                  dihydrochloride, 2,6-dimethylpiperazine, triethylenetetramine
 * org1mass     : [0.0858, 1.3031] grams
 * temp         : [90.0, 180.0]
 * time         : [18.0, 48.0]
 * pH           : [1.0, 7.0]

============================== An example reaction ==============================
 * inorg1       : sodium metavanadate
 * inorg1mass   : 0.156 grams
 * inorg2       : Selenium dioxide
 * inorg2mass   : 1.1275 grams
 * org1         : N,N'-dimethylethylenediamine
 * org1mass     : 0.0938 grams
 * temp         : 90
 * time         : 18.0
 * slowCool     : 1
 * pH           : 1
```

```
Would you like to conduct the experiments for this region? (y/n) : n

Explaining Query with id #945
Certainty 0.016

========== The model is uncertain about the reaction because ... ==============
 * numberInorg >= 3
 * slowCool = True
 * 139.50 < AtomicRadiusMean <= 146.83
 * 352.66 < PearsonElectronegGeom <= 438.68
 * 352.67 < PearsonElectronegMean <= 446.73
 * 54.03 < EAMean <= 121.60

Instances in the batch: 15
================ A potential variation for probing the region ==================
 * inorg1       : Oxovanadium(2+) sulfate, sodium vanadium trioxide,
                  ammonium metavanadate, sodium metavanadate
 * inorg1mass   : [0.0449, 0.3174] grams
 * inorg2       : Potassium Dichromate
 * inorg2mass   : [0.0598, 0.6328] grams
 * inorg3       : Selenium dioxide
 * inorg3mass   : [0.1597, 1.3363] grams
 * org1         : N,N'-dimethylethylenediamine, diethylenetriamine,
                  2-methylpiperazine, 2,5-dimethylpiperazine,
                  1,4-dimethylpiperazine
 * org1mass     : [0.0336, 0.863] grams
 * pH           : [1.0, 5.0]

============================== An example reaction ==============================
 * inorg1       : Oxovanadium(2+) sulfate
 * inorg1mass   : 0.1631 grams
 * inorg2       : Potassium Dichromate
 * inorg2mass   : 0.6328 grams
 * inorg3       : Selenium dioxide
 * inorg3mass   : 0.6427 grams
 * org1         : 2-methylpiperazine
 * org1mass     : 0.1111 grams
 * temp         : 110
 * time         : 24.0
 * slowCool     : 1
 * pH           : 1
```

```
Would you like to conduct the experiments for this region? (y/n) : y

================================================================================
Round 5
================================================================================
Explaining Query with id #1088
Certainty 0.000

========== The model is uncertain about the reaction because ... ==============
 * slowCool = True
 * 251.93 < orgASA_HMax <= 335.55
 * orghbdamsdonGeomAvg <= 0.95
 * 260.83 < orgASA_HArithAvg <= 336.83
 * 257.46 < orgASA_HGeomAvg <= 336.83
 * 256.47 < orgASA_HMin <= 336.83

Instances in the batch: 17
================ A potential variation for probing the region ==================
 * inorg1       : Oxovanadium(2+) sulfate, sodium vanadium trioxide,
                  ammonium metavanadate, vanadium(V) oxide,
                  sodium metavanadate
 * inorg1mass   : [0.0993, 0.3694] grams
 * inorg2       : Selenium dioxide
 * inorg2mass   : [0.361, 1.346] grams
 * org1         : 2,6-dimethylpiperazine, 3-aminoquinuclidine dihydrochloride
 * org1mass     : [0.1207, 0.3982] grams
 * temp         : [90.0, 150.0]
 * time         : [18.0, 48.0]
 * pH           : [2.0, 7.0]

============================== An example reaction ==============================
 * inorg1       : Oxovanadium(2+) sulfate
 * inorg1mass   : 0.3304 grams
 * inorg2       : Selenium dioxide
 * inorg2mass   : 0.5595 grams
 * org1         : 2,6-dimethylpiperazine
 * org1mass     : 0.2062 grams
 * temp         : 90
 * time         : 24.0
 * slowCool     : 1
```

```
 * pH           : 2

Would you like to conduct the experiments for this region? (y/n) : y

Explaining Query with id #1351
Certainty 0.000

=========== The model is uncertain about the reaction because ... ==============
 * PaulingElectronegMean <= 1.84
 * PaulingElectronegGeom <= 1.83
 * slowCool = True
 * PearsonElectronegMean <= 395.05
 * 52.87 < orgASA-GeomAvg <= 74.69

Instances in the batch: 3
================ A potential variation for probing the region ==================
 * inorg1       : Ga2O3, ammonium metavanadate, vanadium(V) oxide
 * inorg1mass   : [0.0642, 0.1646] grams
 * org1         : H3PO3, 2-methylpiperazine
 * org1mass     : [0.0828, 1.0488] grams
 * org2         : 1-methylpiperazine
 * temp         : [110.0, 180.0]
 * time         : [24.0, 96.0]
 * pH           : [2.0, 8.0]

============================== An example reaction ==============================
 * inorg1       : ammonium metavanadate
 * inorg1mass   : 0.1171 grams
 * org1         : 2-methylpiperazine
 * org1mass     : 0.1182 grams
 * temp         : 110
 * time         : 48.0
 * slowCool     : 1
 * pH           : 2

Would you like to conduct the experiments for this region? (y/n) : y

================================================================================
Round 6
================================================================================
Explaining Query with id #102
```

```
Certainty 0.000

========== The model is uncertain about the reaction because ... ==============
 * 1.76 < PaulingElectronegMean <= 2.02
 * orgmaximalprojectionsizeMax <= 5.09
 * 4.98 < orgmaximalprojectionradiusArithAvg <= 5.77
 * slowCool = True
 * orgminimalprojectionradiusGeomAvg <= 3.32

Instances in the batch: 20
================ A potential variation for probing the region ==================
 * inorg1       : Oxovanadium(2+) sulfate, gallium trinitrate,
                  ammonium metavanadate
                  sodium metavanadate
 * inorg1mass   : [0.0301, 1.264] grams
 * inorg2       : sodium tellurite, Potassium Dichromate, tellurium dioxide
 * inorg2mass   : [0.0398, 0.5711] grams
 * inorg3       : Selenium dioxide
 * org1         : N,N'-dimethylethylenediamine
 * org1mass     : [0.0206, 0.7989] grams
 * temp         : [110.0, 180.0]
 * pH           : [1.0, 7.0]

============================== An example reaction ==============================
 * inorg1       : Oxovanadium(2+) sulfate
 * inorg1mass   : 0.0403 grams
 * inorg2       : sodium tellurite
 * inorg2mass   : 0.0605 grams
 * org1         : N,N'-dimethylethylenediamine
 * org1mass     : 0.2248 grams
 * temp         : 110
 * time         : 24.0
 * slowCool     : 1
 * pH           : 3

Would you like to conduct the experiments for this region? (y/n) : y

================================================================================
Round 7
================================================================================
Explaining Query with id #573
```

```
Certainty 0.000

========== The model is uncertain about the reaction because ... ==============
 * 1.76 < PaulingElectronegGeom <= 1.89
 * time > 54.75
 * 1.76 < PaulingElectronegMean <= 1.91
 * orgmaximalprojectionradiusGeomAvg > 5.77
 * 7 <= pH <= 11
 * 1.73 < PaulingElectronegMax <= 2.13

Instances in the batch: 20
================ A potential variation for probing the region ==================
 * inorg1       : sodium tellurite
 * inorg1mass   : [0.0842, 0.2879] grams
 * inorg2       : ammonium metavanadate, sodium metavanadate
 * inorg2mass   : [0.076, 0.1757] grams
 * org1         : 1,5-diaminopentane, 1,6-diaminohexane,
                  1,4-bis(3-aminopropyl)piperazine
 * org1mass     : [0.0719, 0.1285] grams
 * temp         : [75.0, 90.0]
 * time         : [60.0, 84.0]
 * slowCool     : [no, yes]
 * pH           : [8.0, 9.0]

============================== An example reaction ==============================
 * inorg1       : sodium tellurite
 * inorg1mass   : 0.2797 grams
 * inorg2       : sodium metavanadate
 * inorg2mass   : 0.1571 grams
 * org1         : 1,6-diaminohexane
 * org1mass     : 0.0933 grams
 * temp         : 90
 * time         : 60.0
 * slowCool     : 0
 * pH           : 8

Would you like to conduct the experiments for this region? (y/n) : y

================================================================================
Round 8
================================================================================
```

```
Explaining Query with id #32
Certainty 0.000

========== The model is uncertain about the reaction because ... ==============
 * 1.76 < PaulingElectronegMean <= 1.91
 * 1.76 < PaulingElectronegGeom <= 1.89
 * 5 <= pH <= 9
 * 1.85 < PaulingElectronegMax <= 2.13

Instances in the batch: 20
================ A potential variation for probing the region ==================
 * inorg1       : sodium tellurite, vanadium(V) oxide, tellurium dioxide
                  sodium metavanadate
 * inorg1mass   : [0.0295, 10.2772] grams
 * inorg2       : sodium tellurite, ammonium metavanadate, vanadium(V) oxide
                  tellurium dioxide, sodium metavanadate
 * inorg2mass   : [0.0933, 0.761] grams
 * org1         : 1-methylheptane, hexamine, 2-methylpiperazine
                  ethylenediamine, diethylenetriamine, 1,4-diaminobutane
                  1,4-bis(3-aminopropyl)piperazine, 1,3-diaminopropane,
                  tris(2-aminoethyl)amine, (S)-3-aminopiperidine dihydrochloride,
                  S-2-methylpiperazine, 1,4-diazabicyclo[2.2.2]octane
                  N,N,N',N'-Tetramethylethylenediamine, 1,5-diaminopentane,
                  1,6-diaminohexane, piperazine, 1,2-diaminopropane
 * org1mass     : [0.0238, 0.3792] grams
 * org2         : ethanol
 * org2mass     : [3.0, 4.0288] grams
 * oxlike1      : sodium oxalate
 * temp         : [90.0, 180.0]
 * time         : [24.0, 192.0]
 * slowCool     : [no, yes]
 * pH           : [7.0, 9.0]

============================== An example reaction ==============================
 * inorg1       : sodium metavanadate
 * inorg1mass   : 0.0295 grams
 * inorg2       : tellurium dioxide
 * inorg2mass   : 0.4053 grams
 * org1         : diethylenetriamine
 * org1mass     : 0.0238 grams
 * temp         : 110
```

```
 * time         : 24.0
 * slowCool     : 1
 * pH           : 7

Would you like to conduct the experiments for this region? (y/n) : y

================================================================================
Round 9
================================================================================
Explaining Query with id #579
Certainty 0.000

========== The model is uncertain about the reaction because ... ==============
 * 1.84 < PaulingElectronegMean <= 2.02
 * 7 <= pH <= 10
 * 59.75 < time <= 132.00
 * 1.83 < PaulingElectronegGeom <= 1.97
 * 0.04 < EAMeanWeighted <= 0.60

Instances in the batch: 20
================ A potential variation for probing the region ==================
 * inorg1       : sodium tellurite, vanadium(V) oxide, tellurium dioxide
                  sodium metavanadate
 * inorg1mass   : [0.0965, 0.4785] grams
 * inorg2       : sodium tellurite, ammonium metavanadate,
                  potassium metavanadate, sodium metavanadate
 * inorg2mass   : [0.0748, 0.571] grams
 * inorg3       : hydrazine
 * org1         : 1,5-diaminopentane, N,N'-diethylethylenediamine,
                  S-2-methylpiperazine, 1-methylpiperazine,
                  N,N'-dimethylethylenediamine, ethylenediamine,
                  1,3-diaminopropane, piperazine, 1,4-diazabicyclo[2.2.2]octane
                  ethanol, N,N,N',N'-Tetramethylethylenediamine,
                  1,6-diaminohexane, tris(2-aminoethyl)amine, 1,2-diaminopropane
 * org1mass     : [0.0628, 3.0] grams
 * org2         : ethanol
 * org2mass     : [3.95, 4.1005] grams
 * temp         : [75.0, 180.0]
 * time         : [59.5, 120.0]
 * slowCool     : [no, yes]
 * pH           : [8.0, 10.0]
```

```
============================== An example reaction ==============================
 * inorg1       : sodium tellurite
 * inorg1mass   : 0.2715 grams
 * inorg2       : sodium metavanadate
 * inorg2mass   : 0.1431 grams
 * org1         : 1,6-diaminohexane
 * org1mass     : 0.0936 grams
 * temp         : 75
 * time         : 60.0
 * slowCool     : 0
 * pH           : 8

Would you like to conduct the experiments for this region? (y/n) : n

Explaining Query with id #604
Certainty 0.000

=========== The model is uncertain about the reaction because ... ==============
 * 1.84 < PaulingElectronegMean <= 2.02
 * 1.83 < PaulingElectronegGeom <= 1.97
 * 5 <= pH <= 7
 * 59.75 < time <= 132.00
 * numberInorg = 2
 * 0.04 < EAMeanWeighted <= 0.60

Instances in the batch: 3
================ A potential variation for probing the region ==================
 * inorg1       : sodium tellurite, sodium metavanadate
 * inorg1mass   : [0.1495, 0.468] grams
 * inorg2       : sodium tellurite, sodium metavanadate
 * inorg2mass   : [0.1432, 0.2642] grams
 * org1         : 1,6-diaminohexane, 1,2-diaminopropane
 * org1mass     : [0.114, 0.4601] grams
 * temp         : [90.0, 100.0]
 * time         : [60.0, 72.0]
 * slowCool     : [no, yes]
 * pH           : [6.0, 7.0]

============================== An example reaction ==============================
 * inorg1       : sodium metavanadate
```

```
 * inorg1mass   : 0.1495 grams
 * inorg2       : sodium tellurite
 * inorg2mass   : 0.2642 grams
 * org1         : 1,6-diaminohexane
 * org1mass     : 0.136 grams
 * temp         : 90
 * time         : 60.0
 * slowCool     : 1
 * pH           : 6

Would you like to conduct the experiments for this region? (y/n) : y

Explaining Query with id #820
Certainty 0.000

========== The model is uncertain about the reaction because ... ==============
 * numberInorg >= 3
 * 395.05 < PearsonElectronegGeom <= 422.98
 * 1.84 < PaulingElectronegMean <= 2.02
 * 710.52 < IonizationGeom <= 744.48
 * 1.83 < PaulingElectronegGeom <= 1.97
 * 62.22 < orgASA-Min <= 74.99
 * 716.28 < IonizationMean <= 754.18

Instances in the batch: 2
================ A potential variation for probing the region ==================
 * inorg1       : Oxovanadium(2+) sulfate, sodium vanadium trioxide
 * inorg1mass   : [0.1254, 0.3174] grams
 * inorg2       : Potassium Dichromate
 * inorg2mass   : [0.1799, 0.525] grams
 * inorg3       : Selenium dioxide
 * inorg3mass   : [0.5856, 1.3325] grams
 * org1         : 2-methylpiperazine
 * org1mass     : [0.0961, 0.1008] grams

============================== An example reaction ==============================
 * inorg1       : sodium vanadium trioxide
 * inorg1mass   : 0.1254 grams
 * inorg2       : Potassium Dichromate
 * inorg2mass   : 0.1799 grams
 * inorg3       : Selenium dioxide
```

```
 * inorg3mass   : 1.3325 grams
 * org1         : 2-methylpiperazine
 * org1mass     : 0.1008 grams
 * temp         : 110
 * time         : 24.0
 * slowCool     : 1
 * pH           : 2

Would you like to conduct the experiments for this region? (y/n) : y

Explaining Query with id #1147
Certainty 0.000

========== The model is uncertain about the reaction because ... ==============
 * orgminimalprojectionradiusMax <= 2.90
 * orgminimalprojectionradiusGeomAvg <= 2.89
 * orgminimalprojectionradiusArithAvg <= 2.89
 * numberInorg = 2
 * slowCool = True

Instances in the batch: 15
================ A potential variation for probing the region ==================
 * inorg1       : ammonium metavanadate, vanadium(V) oxide, sodium metavanadate
                  tellurium dioxide, sodium fluoride, molybdenum trioxide
 * inorg1mass   : [0.0457, 0.9128] grams
 * inorg2       : HIO3, sodium metavanadate, sodium tellurite
                  Selenium dioxide, tellurium dioxide, Potassium Dichromate
                  ammonium fluoride
 * inorg2mass   : [0.0386, 2.5575] grams
 * org1         : H3PO3, N,N'-dimethylethylenediamine, ethylenediamine
                  1,3-diaminopropane
 * org1mass     : [0.0144, 1.2571] grams
 * org2         : ethylenediamine, ethanol
 * org2mass     : [0.0624, 3.95] grams
 * oxlike1      : sodium oxalate
 * oxlike1mass  : [0.1396, 0.2809] grams
 * temp         : [90.0, 180.0]
 * time         : [17.0, 120.0]
 * pH           : [1.0, 11.0]

============================== An example reaction ==============================
```

```
 * inorg1       : sodium metavanadate
 * inorg1mass   : 0.1278 grams
 * inorg2       : Selenium dioxide
 * inorg2mass   : 1.1174 grams
 * org1         : 1,3-diaminopropane
 * org1mass     : 0.0855 grams
 * oxlike1      : sodium oxalate
 * oxlike1mass  : 0.1396 grams
 * temp         : 90
 * time         : 18.0
 * slowCool     : 1
 * pH           : 2

Would you like to conduct the experiments for this region? (y/n) : y

================================================================================
Round 10
================================================================================
Explaining Query with id #154
Certainty 0.000

========== The model is uncertain about the reaction because ... ==============
 * 1.76 < PaulingElectronegGeom <= 1.89
 * 1.76 < PaulingElectronegMean <= 1.91
 * orgminimalprojectionradiusMax <= 2.84
 * slowCool = True
 * orghbdamsaccGeomAvg <= 0.99
 * numberInorg = 2

Instances in the batch: 20
================ A potential variation for probing the region ==================
 * inorg1       : Oxovanadium(2+) sulfate, sodium tellurite, tellurium dioxide
                  sodium metavanadate
 * inorg1mass   : [0.0292, 0.4345] grams
 * inorg2       : sodium tellurite, ammonium metavanadate, tellurium dioxide
                  sodium metavanadate
 * inorg2mass   : [0.0294, 0.4062] grams
 * org1         : N,N'-dimethylethylenediamine, ethylenediamine
 * org1mass     : [0.0129, 0.2346] grams
 * temp         : [90.0, 110.0]
 * time         : [17.0, 60.0]
```

```
 * pH           : [1.0, 8.0]

============================== An example reaction ==============================
 * inorg1       : sodium tellurite
 * inorg1mass   : 0.1356 grams
 * inorg2       : sodium metavanadate
 * inorg2mass   : 0.0785 grams
 * org1         : ethylenediamine
 * org1mass     : 0.0722 grams
 * temp         : 90
 * time         : 24.0
 * slowCool     : 1
 * pH           : 8

Would you like to conduct the experiments for this region? (y/n) : y

================================================================================
Round 11
================================================================================
Explaining Query with id #608
Certainty 0.000

========== The model is uncertain about the reaction because ... ==============
 * numberInorg = 2
 * 1.76 < PaulingElectronegMean <= 1.87
 * slowCool = True
 * 5.98 < orghbdamsdonMin <= 9.95
 * 5.98 < orghbdamsdonGeomAvg <= 9.95

Instances in the batch: 20
================ A potential variation for probing the region ==================
 * inorg1       : Oxovanadium(2+) sulfate, sodium tellurite, sodium metavanadate
 * inorg1mass   : [0.0319, 0.2463] grams
 * inorg2       : sodium tellurite, ammonium metavanadate, vanadium(V) oxide
                  tellurium dioxide, sodium metavanadate
 * inorg2mass   : [0.0442, 0.4056] grams
 * org1         : 1,6-diaminohexane, diethylenetriamine, 1,4-diaminobutane
                  1,4-bis(3-aminopropyl)piperazine, 1,3-diaminopropane,
                  1-(2-aminoethyl)piperazine, 1,5-diaminopentane,
                  tris(2-aminoethyl)amine
 * org1mass     : [0.0196, 0.2715] grams
```

```
 * oxlike1      : sodium oxalate
 * oxlike1mass  : [0.1034, 0.1068] grams
 * temp         : [90.0, 110.0]
 * time         : [24.0, 84.0]
 * pH           : [1.0, 8.0]

============================== An example reaction ==============================
 * inorg1       : sodium tellurite
 * inorg1mass   : 0.1457 grams
 * inorg2       : sodium metavanadate
 * inorg2mass   : 0.0745 grams
 * org1         : 1,6-diaminohexane
 * org1mass     : 0.1391 grams
 * temp         : 90
 * time         : 24.0
 * slowCool     : 1
 * pH           : 8

Would you like to conduct the experiments for this region? (y/n) : y

================================================================================
Round 12
================================================================================
Explaining Query with id #146
Certainty 0.000

========== The model is uncertain about the reaction because ... ==============
 * 5.14 < orghbdamsdonGeomAvg <= 5.80
 * 7.41 < orgavgpol_pHdependentArithAvg <= 7.47
 * orgminimalprojectionradiusGeomAvg <= 2.63
 * numberInorg = 2
 * orgminimalprojectionradiusArithAvg <= 2.63
 * slowCool = True

Instances in the batch: 7
================ A potential variation for probing the region ==================
 * inorg1       : sodium tellurite
 * inorg1mass   : [0.1396, 0.3175] grams
 * inorg2       : sodium metavanadate
 * inorg2mass   : [0.0703, 0.1728] grams
 * org1         : ethylenediamine
```

```
 * org1mass     : [0.074, 0.1166] grams
 * time         : [24.0, 60.0]

============================== An example reaction ==============================
 * inorg1       : sodium tellurite
 * inorg1mass   : 0.2562 grams
 * inorg2       : sodium metavanadate
 * inorg2mass   : 0.1557 grams
 * org1         : ethylenediamine
 * org1mass     : 0.1119 grams
 * temp         : 90
 * time         : 60.0
 * slowCool     : 1
 * pH           : 8

Would you like to conduct the experiments for this region? (y/n) : y

Explaining Query with id #902
Certainty 0.000

=========== The model is uncertain about the reaction because ... ==============
 * numberInorg = 2
 * numberComponents >= 4
 * 5.98 < orghbdamsdonMin <= 9.95
 * slowCool = True
 * 5.98 < orghbdamsdonGeomAvg <= 9.95

Instances in the batch: 13
================ A potential variation for probing the region ==================
 * inorg1       : ammonium metavanadate, sodium metavanadate
 * inorg1mass   : [0.1159, 0.2339] grams
 * inorg2       : Selenium dioxide
 * inorg2mass   : [0.4465, 1.3527] grams
 * org1         : diethylenetriamine, 1,4-diaminobutane, ethylenediamine
                  1,3-diaminopropane, 1,2-diaminocyclohexane, 1,5-diaminopentane
 * org1mass     : [0.0653, 0.1916] grams
 * oxlike1      : sodium oxalate
 * oxlike1mass  : [0.1014, 0.2774] grams
 * temp         : [90.0, 110.0]
 * time         : [16.0, 60.0]
 * pH           : [1.0, 2.0]
```

```
============================== An example reaction ==============================
 * inorg1       : sodium metavanadate
 * inorg1mass   : 0.1269 grams
 * inorg2       : Selenium dioxide
 * inorg2mass   : 1.0123 grams
 * org1         : 1,4-diaminobutane
 * org1mass     : 0.09 grams
 * oxlike1      : sodium oxalate
 * oxlike1mass  : 0.1726 grams
 * temp         : 90
 * time         : 18.0
 * slowCool     : 1
 * pH           : 2

Would you like to conduct the experiments for this region? (y/n) : y

================================================================================
Round 13
================================================================================
Explaining Query with id #401
Certainty 0.000

=========== The model is uncertain about the reaction because ... ==============
 * AtomicRadiusGeomWeighted > 1.08
 * hardnessMinWeighted > 1.91
 * AtomicRadiusMeanWeighted > 1.17
 * IonizationMinWeighted > 4.20
 * PearsonElectronegGeom <= 352.66

Instances in the batch: 19
================ A potential variation for probing the region ==================
 * inorg1       : Oxovanadium(2+) sulfate, sodium vanadium trioxide,
                  potassium vanadium trioxide
 * inorg1mass   : [0.9315, 1.756] grams
 * org1         : H3PO3, N,N'-dimethylethylenediamine, ethylenediamine
                  1,3-diaminopropane, imidazole
 * org1mass     : [0.0104, 0.5076] grams
 * org2         : 1,6-diaminohexane, 1,3-diaminopropane
 * org2mass     : [0.5995, 0.9218] grams
 * temp         : [90.0, 110.0]
```

```
 * time         : [24.0, 48.0]
 * pH           : [1.0, 7.0]

============================== An example reaction ==============================
 * inorg1       : Oxovanadium(2+) sulfate
 * inorg1mass   : 1.138 grams
 * org1         : H3PO3
 * org1mass     : 0.5076 grams
 * org2         : 1,3-diaminopropane
 * org2mass     : 0.5995 grams
 * temp         : 90
 * time         : 48.0
 * slowCool     : 1
 * pH           : 6

Would you like to conduct the experiments for this region? (y/n) : y

Explaining Query with id #2105
Certainty 0.000

=========== The model is uncertain about the reaction because ... ==============
 * numberInorg = 2
 * slowCool = True
 * 5.98 < orghbdamsdonMin <= 7.99
 * orghbdamsaccGeomAvg <= 0.99

Instances in the batch: 1
================ A potential variation for probing the region ==================
 * inorg1       : ammonium metavanadate
 * inorg2       : Selenium dioxide
 * org1         : 1,5-diaminopentane

============================== An example reaction ==============================
 * inorg1       : ammonium metavanadate
 * inorg1mass   : 0.1663 grams
 * inorg2       : Selenium dioxide
 * inorg2mass   : 0.0715 grams
 * org1         : 1,5-diaminopentane
 * org1mass     : 0.1448 grams
 * temp         : 90
 * time         : 19.0
```

```
 * slowCool     : 1
 * pH           : 7

Would you like to conduct the experiments for this region? (y/n) : y

================================================================================
Round 14
================================================================================
Explaining Query with id #2
Certainty 0.000

========== The model is uncertain about the reaction because ... ==============
 * AtomicRadiusMeanWeighted > 1.17
 * slowCool = True
 * numberInorg = 2
 * temp >= 113

Instances in the batch: 2
================ A potential variation for probing the region ==================
 * inorg1       : Oxovanadium(2+) sulfate, sodium metavanadate
 * inorg1mass   : [0.1325, 0.1801] grams
 * inorg2       : Potassium Dichromate
 * inorg2mass   : [2.006, 2.4882] grams
 * org1         : 1-methylpiperazine, diethylenetriamine
 * org1mass     : [0.1047, 0.1053] grams

============================== An example reaction ==============================
 * inorg1       : Oxovanadium(2+) sulfate
 * inorg1mass   : 0.1801 grams
 * inorg2       : Potassium Dichromate
 * inorg2mass   : 2.006 grams
 * org1         : 1-methylpiperazine
 * org1mass     : 0.1053 grams
 * temp         : 130
 * time         : 24.0
 * slowCool     : 1
 * pH           : 2

Would you like to conduct the experiments for this region? (y/n) : n

Explaining Query with id #174
```

```
Certainty 0.000

========== The model is uncertain about the reaction because ... ==============
 * slowCool = True
 * PaulingElectronegGeom <= 1.64
 * PaulingElectronegMean <= 1.64
 * 14.45 < EAGeom <= 53.82
 * 14.45 < EAMean <= 54.03

Instances in the batch: 20
================ A potential variation for probing the region ==================
 * inorg1       : Oxovanadium(2+) sulfate, ammonium metavanadate,
                  sodium metavanadate, vanadium(V) oxide,
                  potassium vanadium trioxide
 * inorg1mass   : [0.0509, 0.8439] grams
 * inorg2       : ammonium fluoride
 * org1         : H3PO3, diethylenetriamine, ethylenediamine,
                  1,3-diaminopropane, 3-aminopiperidine dihydrochloride,
                  1-methylpiperazine
 * org1mass     : [0.0468, 1.0654] grams
 * org2         : diethylenetriamine, ethylenediamine, 1,3-diaminopropane,
                  3-aminopyrrolidine dihydrochloride, triethylenetetramine,
                  N,N,N',N'-Tetramethylethylenediamine, 1,6-diaminohexane,
                  1-methylpiperazine, 1,2-diaminopropane
 * org2mass     : [0.0157, 0.7071] grams
 * temp         : [110.0, 180.0]
 * time         : [20.0, 120.0]
 * pH           : [1.0, 8.0]

============================== An example reaction ==============================
 * inorg1       : vanadium(V) oxide
 * inorg1mass   : 0.1853 grams
 * org1         : H3PO3
 * org1mass     : 0.5574 grams
 * org2         : 1,3-diaminopropane
 * org2mass     : 0.3635 grams
 * temp         : 170
 * time         : 120.0
 * slowCool     : 1
 * pH           : 6
```

```
Would you like to conduct the experiments for this region? (y/n) : n

Explaining Query with id #1158
Certainty 0.000

========== The model is uncertain about the reaction because ... ==============
 * slowCool = True
 * numberInorg = 2
 * numberComponents >= 4
 * 106.66 < orgASA_HGeomAvg <= 142.58

Instances in the batch: 20
================ A potential variation for probing the region ==================
 * inorg1       : sodium metavanadate
 * inorg1mass   : [0.0978, 0.2] grams
 * inorg2       : sodium tellurite, Selenium dioxide
 * inorg2mass   : [0.349, 1.3553] grams
 * org1         : (S)-3-aminopyrrolidine, 1,3-diaminopropane,
                  3-aminopyrrolidine dihydrochloride
                  R-3-aminopyrrolidine dihydrochloride,
                  1,3,5,7- tetraazaadamantane, 3-aminopyrrolidine,
                  imidazole
 * org1mass     : [0.072, 0.1917] grams
 * oxlike1      : sodium oxalate
 * oxlike1mass  : [0.1019, 0.3497] grams
 * temp         : [90.0, 130.0]
 * time         : [16.0, 60.0]
 * pH           : [1.0, 7.0]

============================== An example reaction ==============================
 * inorg1       : sodium metavanadate
 * inorg1mass   : 0.0978 grams
 * inorg2       : Selenium dioxide
 * inorg2mass   : 1.1168 grams
 * org1         : 1,3-diaminopropane
 * org1mass     : 0.0725 grams
 * oxlike1      : sodium oxalate
 * oxlike1mass  : 0.2144 grams
 * temp         : 90
 * time         : 16.0
 * slowCool     : 1
```

```
 * pH           : 1

Would you like to conduct the experiments for this region? (y/n) : y

================================================================================
Round 15
================================================================================
Explaining Query with id #88
Certainty 0.000

=========== The model is uncertain about the reaction because ... ==============
 * orgASA-Max > 97.12
 * orghbdamsaccGeomAvg <= 0.13
 * 40.01 < orgASA-ArithAvg <= 87.19
 * orgminimalprojectionradiusGeomAvg <= 2.92

Instances in the batch: 9
================ A potential variation for probing the region ==================
 * inorg1       : vanadium(V) oxide, sodium metavanadate
 * inorg1mass   : [0.0457, 0.1884] grams
 * inorg2       : sodium tellurite, ammonium fluoride
 * inorg2mass   : [0.0421, 0.1152] grams
 * org1         : H3PO3, 1-methylheptane
 * org1mass     : [0.1241, 1.0235] grams
 * org2         : diethylenetriamine, 1,4-diaminobutane, ethylenediamine
                  1,3-diaminopropane, ethanol
 * org2mass     : [0.107, 4.0539] grams
 * temp         : [110.0, 170.0]
 * time         : [48.0, 120.0]
 * slowCool     : [no, yes]
 * pH           : [1.0, 12.0]

============================== An example reaction ==============================
 * inorg1       : vanadium(V) oxide
 * inorg1mass   : 0.1846 grams
 * inorg2       : sodium tellurite
 * inorg2mass   : 0.1152 grams
 * org1         : 1-methylheptane
 * org1mass     : 0.1426 grams
 * org2         : ethanol
 * org2mass     : 4.017 grams
```

```
 * temp         : 110
 * time         : 72.0
 * slowCool     : 0
 * pH           : 12

Would you like to conduct the experiments for this region? (y/n) : n

Explaining Query with id #2047
Certainty 0.017

========== The model is uncertain about the reaction because ... ==============
 * slowCool = True
 * 1.64 < PaulingElectronegGeom <= 1.65
 * 352.67 < PearsonElectronegMean <= 366.38
 * 53.82 < EAGeom <= 67.80
 * 1.64 < PaulingElectronegMax <= 1.73
 * 1.64 < PaulingElectronegMean <= 1.65
 * 651.90 < IonizationMax <= 668.60

Instances in the batch: 20
================ A potential variation for probing the region ==================
 * inorg1       : Oxovanadium(2+) sulfate, ammonium metavanadate,
                  sodium metavanadate
 * inorg1mass   : [0.1098, 0.3433] grams
 * inorg2       : Potassium Dichromate
 * inorg2mass   : [0.5152, 2.4882] grams
 * org1         : 2-methylpiperazine, 1-methylpiperazine,
                  N,N'-dimethylethylenediamine,
                  3-aminopyrrolidine dihydrochloride,
                  1,4-dimethylpiperazine, diethylenetriamine,
                  2,5-dimethylpyrazine
 * org1mass     : [0.0853, 0.1661] grams
 * oxlike1      : sodium oxalate
 * oxlike1mass  : [0.0987, 0.1257] grams
 * temp         : [110.0, 180.0]
 * time         : [24.0, 89.0]
 * pH           : [1.0, 2.0]

============================== An example reaction ==============================
 * inorg1       : Oxovanadium(2+) sulfate
 * inorg1mass   : 0.1851 grams
```

```
 * inorg2       : Potassium Dichromate
 * inorg2mass   : 1.0075 grams
 * org1         : 1,4-dimethylpiperazine
 * org1mass     : 0.1217 grams
 * temp         : 110
 * time         : 24.0
 * slowCool     : 1
 * pH           : 1

Would you like to conduct the experiments for this region? (y/n) : y

================================================================================
Round 16
================================================================================
Explaining Query with id #1653
Certainty 0.000

========== The model is uncertain about the reaction because ... ==============
 * slowCool = True
 * 42.93 < orgASA-GeomAvg <= 87.19
 * 42.93 < orgASA-Min <= 87.19
 * 42.93 < orgASA-ArithAvg <= 87.19

Instances in the batch: 20
================ A potential variation for probing the region ==================
 * inorg1       : Oxovanadium(2+) sulfate, gallium trinitrate,
                  sodium vanadium trioxide, ammonium metavanadate, ZnO,
                  sodium metavanadate, sodium tellurite,
                  potassium vanadium trioxide, sodium fluoride,
                  molybdenum trioxide
 * inorg1mass   : [0.0328, 1.638] grams
 * inorg2       : HIO3, sodium metavanadate, sodium tellurite
                  Selenium dioxide, tellurium dioxide, Potassium Dichromate
                  ammonium fluoride
 * inorg2mass   : [0.0299, 2.2307] grams
 * inorg3       : Selenium dioxide
 * org1         : 2-methylpiperazine, N,N'-dimethylethylenediamine,
                  1,4-bis(3-aminopropyl)piperazine, 1,3-diaminopropane,
                  2-methylpyrazine, (S)-3-aminoquinuclidine dihydrochloride,
                  1,2-diaminocyclohexane, 1,5-diaminopentane
 * org1mass     : [0.0153, 2.3417] grams
```

```
 * oxlike1      : sodium oxalate
 * oxlike1mass  : [0.1022, 0.1404] grams
 * temp         : [80.0, 180.0]
 * time         : [15.0, 96.0]
 * pH           : [1.0, 13.0]

============================== An example reaction ==============================
 * inorg1       : sodium metavanadate
 * inorg1mass   : 0.1302 grams
 * inorg2       : Selenium dioxide
 * inorg2mass   : 1.034 grams
 * org1         : 1,4-bis(3-aminopropyl)piperazine
 * org1mass     : 0.2546 grams
 * temp         : 90
 * time         : 15.0
 * slowCool     : 1
 * pH           : 1

Would you like to conduct the experiments for this region? (y/n) : n

Explaining Query with id #2196
Certainty 0.000

=========== The model is uncertain about the reaction because ... ==============
 * 32.64 < orgmaximalprojectionareaMin <= 32.87
 * 150.47 < orgASA+Min <= 195.40
 * 36.80 < oxlikepolarsurfaceareaMax <= 76.01

Instances in the batch: 2
================ A potential variation for probing the region ==================
 * inorg1       : sodium metavanadate
 * inorg1mass   : [0.1214, 0.1279] grams
 * inorg2       : Selenium dioxide
 * inorg2mass   : [0.6659, 0.6962] grams
 * org1         : 3-aminopyrrolidine dihydrochloride
 * org1mass     : [0.1552, 0.1574] grams
 * oxlike1      : sodium oxalate
 * oxlike1mass  : [0.1171, 0.1435] grams

============================== An example reaction ==============================
 * inorg1       : sodium metavanadate
```

```
 * inorg1mass   : 0.1214 grams
 * inorg2       : Selenium dioxide
 * inorg2mass   : 0.6659 grams
 * org1         : 3-aminopyrrolidine dihydrochloride
 * org1mass     : 0.1574 grams
 * oxlike1      : sodium oxalate
 * oxlike1mass  : 0.1171 grams
 * temp         : 90
 * time         : 24.0
 * slowCool     : 0
 * pH           : 1

Would you like to conduct the experiments for this region? (y/n) : y

Explaining Query with id #914
Certainty 0.013

========== The model is uncertain about the reaction because ... ==============
 * 42.93 < orgASA-Min <= 87.19
 * 42.93 < orgASA-ArithAvg <= 87.19
 * 1.76 < PaulingElectronegMean <= 1.87
 * 42.93 < orgASA-GeomAvg <= 87.19
 * orgminimalprojectionradiusMin <= 2.93

Instances in the batch: 18
================ A potential variation for probing the region ==================
 * inorg1       : Oxovanadium(2+) sulfate, gallium trinitrate, vanadium(V) oxide
                  sodium metavanadate, sodium tellurite, tellurium dioxide
 * inorg1mass   : [0.0328, 0.4774] grams
 * inorg2       : sodium tellurite, tellurium dioxide, sodium metavanadate
 * inorg2mass   : [0.0565, 0.4486] grams
 * org1         : 1-methylheptane, N,N'-diethylethylenediamine,
                  N,N'-dimethylethylenediamine, 1,3-diaminopropane,
                  2-aminooctane, 1,6-diaminohexane,
                  N,N,N',N'-Tetramethylethylenediamine,
                  1-ethylpiperazine, 1,5-diaminopentane
 * org1mass     : [0.0153, 0.7683] grams
 * org2         : ethanol
 * org2mass     : [1.9788, 4.017] grams
 * temp         : [90.0, 180.0]
 * time         : [24.0, 192.0]
```

```
 * slowCool     : [no, yes]
 * pH           : [3.0, 12.0]

============================== An example reaction ==============================
 * inorg1       : sodium tellurite
 * inorg1mass   : 0.2767 grams
 * inorg2       : sodium metavanadate
 * inorg2mass   : 0.0756 grams
 * org1         : 1,3-diaminopropane
 * org1mass     : 0.0708 grams
 * temp         : 90
 * time         : 60.0
 * slowCool     : 0
 * pH           : 8

Would you like to conduct the experiments for this region? (y/n) : y

================================================================================
Round 17
================================================================================
Explaining Query with id #413
Certainty 0.000

=========== The model is uncertain about the reaction because ... ==============
 * slowCool = True
 * 12 <= time <= 16
 * 13.34 < orgavgpol_pHdependentMax <= 13.44
 * 105.32 < oxlikevanderwaalsGeomAvg <= 215.85

Instances in the batch: 9
================ A potential variation for probing the region ==================
 * inorg1       : sodium metavanadate
 * inorg1mass   : [0.0989, 0.1486] grams
 * inorg2       : Selenium dioxide
 * inorg2mass   : [0.8939, 1.3356] grams
 * org1         : 1,4-diazabicyclo[2.2.2]octane
 * org1mass     : [0.134, 0.2478] grams
 * oxlike1      : sodium oxalate
 * oxlike1mass  : [0.225, 0.4166] grams

============================== An example reaction ==============================
```

```
 * inorg1       : sodium metavanadate
 * inorg1mass   : 0.124 grams
 * inorg2       : Selenium dioxide
 * inorg2mass   : 1.139 grams
 * org1         : 1,4-diazabicyclo[2.2.2]octane
 * org1mass     : 0.134 grams
 * oxlike1      : sodium oxalate
 * oxlike1mass  : 0.225 grams
 * temp         : 90
 * time         : 16.0
 * slowCool     : 1
 * pH           : 1

Would you like to conduct the experiments for this region? (y/n) : y

Explaining Query with id #421
Certainty 0.000

========== The model is uncertain about the reaction because ... ==============
 * slowCool = True
 * oxlikeASA+GeomAvg > 51.28
 * oxlikeASA+Max > 51.28
 * 211.05 < orgASAArithAvg <= 212.28

Instances in the batch: 11
================ A potential variation for probing the region ==================
 * inorg1       : sodium metavanadate
 * inorg1mass   : [0.0989, 0.1486] grams
 * inorg2       : Selenium dioxide
 * inorg2mass   : [0.8963, 1.3361] grams
 * org1         : 1,4-diazabicyclo[2.2.2]octane
 * org1mass     : [0.134, 0.209] grams
 * oxlike1      : sodium oxalate
 * oxlike1mass  : [0.139, 0.3474] grams
 * time         : [16.0, 18.0]

============================== An example reaction ==============================
 * inorg1       : sodium metavanadate
 * inorg1mass   : 0.126 grams
 * inorg2       : Selenium dioxide
 * inorg2mass   : 1.133 grams
```

```
 * org1         : 1,4-diazabicyclo[2.2.2]octane
 * org1mass     : 0.222 grams
 * oxlike1      : sodium oxalate
 * oxlike1mass  : 0.169 grams
 * temp         : 90
 * time         : 18.0
 * slowCool     : 1
 * pH           : 1

Would you like to conduct the experiments for this region? (y/n) : y

Explaining Query with id #1150
Certainty 0.000

========== The model is uncertain about the reaction because ... ==============
 * slowCool = True
 * oxlikeASA+ArithAvg > 51.28
 * oxlikeASA+GeomAvg > 51.28
 * 42.93 < orgASA-ArithAvg <= 87.19

Instances in the batch: 3
================ A potential variation for probing the region ==================
 * inorg1       : sodium metavanadate
 * inorg1mass   : [0.0989, 0.1304] grams
 * inorg2       : Selenium dioxide
 * inorg2mass   : [0.8979, 1.1273] grams
 * org1         : 1,3-diaminopropane
 * org1mass     : [0.0787, 0.0927] grams
 * oxlike1      : sodium oxalate
 * oxlike1mass  : [0.1404, 0.3477] grams
 * time         : [17.0, 18.0]

============================== An example reaction ==============================
 * inorg1       : sodium metavanadate
 * inorg1mass   : 0.0989 grams
 * inorg2       : Selenium dioxide
 * inorg2mass   : 1.1273 grams
 * org1         : 1,3-diaminopropane
 * org1mass     : 0.0881 grams
 * oxlike1      : sodium oxalate
 * oxlike1mass  : 0.1404 grams
```

```
 * temp         : 90
 * time         : 18.0
 * slowCool     : 1
 * pH           : 1

Would you like to conduct the experiments for this region? (y/n) : y

================================================================================
Round 18
================================================================================
Explaining Query with id #1174
Certainty 0.000

=========== The model is uncertain about the reaction because ... ==============
 * slowCool = True
 * oxlikeASA+GeomAvg > 51.28
 * orgmolpolMin <= 10.01
 * orgminimalprojectionradiusGeomAvg <= 3.11

Instances in the batch: 20
================ A potential variation for probing the region ==================
 * inorg1       : sodium metavanadate
 * inorg1mass   : [0.0996, 0.1494] grams
 * inorg2       : Selenium dioxide
 * inorg2mass   : [1.111, 1.3364] grams
 * org1         : 1,3-diaminopropane
 * org1mass     : [0.0569, 0.1115] grams
 * oxlike1      : sodium oxalate
 * oxlike1mass  : [0.112, 0.3508] grams
 * time         : [16.0, 19.0]

============================== An example reaction ==============================
 * inorg1       : sodium metavanadate
 * inorg1mass   : 0.0996 grams
 * inorg2       : Selenium dioxide
 * inorg2mass   : 1.1121 grams
 * org1         : 1,3-diaminopropane
 * org1mass     : 0.0818 grams
 * oxlike1      : sodium oxalate
 * oxlike1mass  : 0.2778 grams
 * temp         : 90
```

```
 * time         : 17.0
 * slowCool     : 1
 * pH           : 1

Would you like to conduct the experiments for this region? (y/n) : y

================================================================================
Round 19
================================================================================
Explaining Query with id #1256
Certainty 0.000

========== The model is uncertain about the reaction because ... ==============
 * time > 54.75
 * 1.76 < PaulingElectronegGeom <= 1.89
 * 1.76 < PaulingElectronegMean <= 1.91
 * 0.13 < PearsonElectronegMeanWeighted <= 1.36
 * 0.11 < hardnessMaxWeighted <= 2.04

Instances in the batch: 20
================ A potential variation for probing the region ==================
 * inorg1       : sodium tellurite, vanadium(V) oxide, tellurium dioxide
                  sodium metavanadate
 * inorg1mass   : [0.1208, 0.4801] grams
 * inorg2       : sodium tellurite, ammonium metavanadate,
                  potassium metavanadate, tellurium dioxide, sodium metavanadate
 * inorg2mass   : [0.0768, 0.1709] grams
 * inorg3       : hydrazine
 * org1         : 1,5-diaminopentane, N,N'-diethylethylenediamine,
                  S-2-methylpiperazine, 1-methylpiperazine, 1-methylheptane,
                  ethylenediamine, 1,3-diaminopropane, 2-aminooctane,
                  piperazine, ethanol, N,N,N',N'-Tetramethylethylenediamine,
                  1-ethylpiperazine, 1,6-diaminohexane, tris(2-aminoethyl)amine,
                  1,2-diaminopropane
 * org1mass     : [0.0628, 3.0] grams
 * org2         : ethanol
 * org2mass     : [3.9414, 4.1169] grams
 * temp         : [90.0, 180.0]
 * time         : [59.5, 192.0]
 * slowCool     : [no, yes]
 * pH           : [3.0, 12.0]
```

```
============================== An example reaction ==============================
 * inorg1       : tellurium dioxide
 * inorg1mass   : 0.4755 grams
 * inorg2       : sodium metavanadate
 * inorg2mass   : 0.1592 grams
 * org1         : ethylenediamine
 * org1mass     : 0.3828 grams
 * temp         : 90
 * time         : 192.0
 * slowCool     : 0
 * pH           : 8

Would you like to conduct the experiments for this region? (y/n) : n

Explaining Query with id #2192
Certainty 0.000

=========== The model is uncertain about the reaction because ... ==============
 * 7.36 < IonizationMaxWeighted <= 7.57
 * 29.45 < oxlikeASA+GeomAvg <= 51.28
 * 29.45 < oxlikeASA+ArithAvg <= 51.28
 * orgmolpolMax <= 14.75
 * 18.44 < orgrefractivityMin <= 25.33

Instances in the batch: 1
================ A potential variation for probing the region ==================
 * inorg1       : sodium metavanadate
 * inorg2       : Selenium dioxide
 * org1         : 3-aminopyrrolidine dihydrochloride
 * oxlike1      : sodium oxalate

============================== An example reaction ==============================
 * inorg1       : sodium metavanadate
 * inorg1mass   : 0.128 grams
 * inorg2       : Selenium dioxide
 * inorg2mass   : 0.8896 grams
 * org1         : 3-aminopyrrolidine dihydrochloride
 * org1mass     : 0.151 grams
 * oxlike1      : sodium oxalate
 * oxlike1mass  : 0.1725 grams
```

```
 * temp         : 90
 * time         : 24.0
 * slowCool     : 0
 * pH           : 2

Would you like to conduct the experiments for this region? (y/n) : y

Explaining Query with id #1719
Certainty 0.019

=========== The model is uncertain about the reaction because ... ==============
 * orgASA-Max > 97.12
 * PaulingElectronegGeom <= 1.64
 * PaulingElectronegMean <= 1.64
 * 14.45 < EAMean <= 54.03
 * 14.45 < EAGeom <= 53.82
 * slowCool = True

Instances in the batch: 19
================ A potential variation for probing the region ==================
 * inorg1       : Oxovanadium(2+) sulfate, vanadium(V) oxide, sodium metavanadate
 * inorg1mass   : [0.0457, 0.8439] grams
 * inorg2       : ammonium fluoride
 * inorg2mass   : [0.0358, 0.0421] grams
 * org1         : H3PO3
 * org1mass     : [0.0468, 1.0654] grams
 * org2         : 1,6-diaminohexane, 3-aminoquinuclidine dihydrochloride,
                  diethylenetriamine, ethylenediamine, 1,3-diaminopropane,
                  3-aminopyrrolidine dihydrochloride, triethylenetetramine,
                  N,N,N',N'-Tetramethylethylenediamine, 1,5-diaminopentane,
                  1-methylpiperazine, 1,2-diaminopropane
 * org2mass     : [0.0157, 0.7071] grams
 * temp         : [110.0, 180.0]
 * time         : [24.0, 120.0]
 * pH           : [1.0, 8.0]

============================== An example reaction ==============================
 * inorg1       : vanadium(V) oxide
 * inorg1mass   : 0.1016 grams
 * org1         : H3PO3
 * org1mass     : 0.4158 grams
```

```
 * org2         : ethylenediamine
 * org2mass     : 0.2817 grams
 * temp         : 170
 * time         : 96.0
 * slowCool     : 1
 * pH           : 6

Would you like to conduct the experiments for this region? (y/n) : n

Explaining Query with id #2234
Certainty 0.022

========== The model is uncertain about the reaction because ... ==============
 * orgpolarsurfaceareaMin > 91.96
 * orgASA-GeomAvg > 105.87
 * orgASA-ArithAvg > 105.69
 * orgpolarsurfaceareaGeomAvg > 91.96
 * orgASA-Min > 93.46
 * orgASA-Max > 97.12

Instances in the batch: 5
================ A potential variation for probing the region ==================
 * inorg1       : molybdenum trioxide
 * inorg1mass   : [0.0837, 0.4915] grams
 * org1         : Lysine
 * org1mass     : [0.0869, 0.5071] grams

============================== An example reaction ==============================
 * inorg1       : molybdenum trioxide
 * inorg1mass   : 0.4915 grams
 * org1         : Lysine
 * org1mass     : 0.0904 grams
 * temp         : 180
 * time         : 48.0
 * slowCool     : 0
 * pH           : 5

Would you like to conduct the experiments for this region? (y/n) : n

Explaining Query with id #2119
Certainty 0.023
```

```
========== The model is uncertain about the reaction because ... ==============
 * slowCool = True
 * orgmaximalprojectionradiusArithAvg > 5.77
 * 42.93 < orgASA-Min <= 87.19
 * 42.93 < orgASA-Max <= 87.19
 * 42.93 < orgASA-GeomAvg <= 87.19
 * orgminimalprojectionsizeGeomAvg > 11.28

Instances in the batch: 19
================ A potential variation for probing the region ==================
 * inorg1       : Oxovanadium(2+) sulfate, sodium tellurite,
                  ammonium metavanadate, sodium metavanadate
 * inorg1mass   : [0.0239, 0.417] grams
 * inorg2       : Selenium dioxide, sodium metavanadate
 * inorg2mass   : [0.0479, 1.1872] grams
 * org1         : 1,5-diaminopentane, 1,4-bis(3-aminopropyl)piperazine,
                  (S)-3-aminoquinuclidine dihydrochloride
                  1,6-diaminohexane
 * org1mass     : [0.0322, 0.824] grams
 * temp         : [90.0, 150.0]
 * time         : [4.0, 96.0]
 * pH           : [1.0, 8.0]

============================== An example reaction ==============================
 * inorg1       : ammonium metavanadate
 * inorg1mass   : 0.3128 grams
 * inorg2       : Selenium dioxide
 * inorg2mass   : 0.0571 grams
 * org1         : 1,5-diaminopentane
 * org1mass     : 0.0322 grams
 * temp         : 90
 * time         : 21.0
 * slowCool     : 1
 * pH           : 3

Would you like to conduct the experiments for this region? (y/n) : y

================================================================================
Round 20
================================================================================
```

```
Explaining Query with id #1650
Certainty 0.000

========== The model is uncertain about the reaction because ... ==============
 * slowCool = True
 * orgmaximalprojectionsizeGeomAvg > 6.95
 * 386.08 < orgvanderwaalsMax <= 404.04
 * orgmaximalprojectionradiusGeomAvg > 6.38

Instances in the batch: 15
================ A potential variation for probing the region ==================
 * inorg1       : sodium metavanadate
 * inorg1mass   : [0.1195, 0.1582] grams
 * inorg2       : Selenium dioxide
 * inorg2mass   : [0.6614, 1.2624] grams
 * org1         : (S)-3-aminoquinuclidine dihydrochloride,
                  1,4-bis(3-aminopropyl)piperazine
 * org1mass     : [0.1652, 0.2805] grams
 * temp         : [90.0, 150.0]
 * time         : [4.0, 20.0]

============================== An example reaction ==============================
 * inorg1       : sodium metavanadate
 * inorg1mass   : 0.1213 grams
 * inorg2       : Selenium dioxide
 * inorg2mass   : 1.1078 grams
 * org1         : 1,4-bis(3-aminopropyl)piperazine
 * org1mass     : 0.2627 grams
 * temp         : 90
 * time         : 6.0
 * slowCool     : 1
 * pH           : 1

Would you like to conduct the experiments for this region? (y/n) : y

Explaining Query with id #2234
Certainty 0.019

========== The model is uncertain about the reaction because ... ==============
 * orgASA-ArithAvg > 90.38
 * orgASA-GeomAvg > 102.13
```

```
 * orgASA-Min > 87.19
 * orgASA-Max > 96.26
 * 366.36 < PearsonElectronegGeom <= 395.05

Instances in the batch: 5
================ A potential variation for probing the region ==================
 * inorg1       : molybdenum trioxide
 * inorg1mass   : [0.0837, 0.4915] grams
 * org1         : Lysine
 * org1mass     : [0.0869, 0.5071] grams

============================== An example reaction ==============================
 * inorg1       : molybdenum trioxide
 * inorg1mass   : 0.4915 grams
 * org1         : Lysine
 * org1mass     : 0.0904 grams
 * temp         : 180
 * time         : 48.0
 * slowCool     : 0
 * pH           : 5

Would you like to conduct the experiments for this region? (y/n) : n

Explaining Query with id #455
Certainty 0.026

========== The model is uncertain about the reaction because ... ==============
 * slowCool = True
 * 0.06 < inorg-water-moleratio <= 0.07
 * oxlikeASA+Min > 51.28
 * oxlikeASA+GeomAvg > 51.28

Instances in the batch: 5
================ A potential variation for probing the region ==================
 * inorg1       : sodium metavanadate
 * inorg1mass   : [0.0989, 0.149] grams
 * inorg2       : Selenium dioxide
 * inorg2mass   : [1.1137, 1.1224] grams
 * org1         : 1,4-diazabicyclo[2.2.2]octane, 1,3-diaminopropane
 * org1mass     : [0.0932, 0.2461] grams
 * oxlike1      : sodium oxalate
```

```
 * oxlike1mass  : [0.2785, 0.3503] grams

============================== An example reaction ==============================
 * inorg1       : sodium metavanadate
 * inorg1mass   : 0.0989 grams
 * inorg2       : Selenium dioxide
 * inorg2mass   : 1.1224 grams
 * org1         : 1,4-diazabicyclo[2.2.2]octane
 * org1mass     : 0.2046 grams
 * oxlike1      : sodium oxalate
 * oxlike1mass  : 0.2798 grams
 * temp         : 90
 * time         : 18.0
 * slowCool     : 1
 * pH           : 1

Would you like to conduct the experiments for this region? (y/n) : y
```