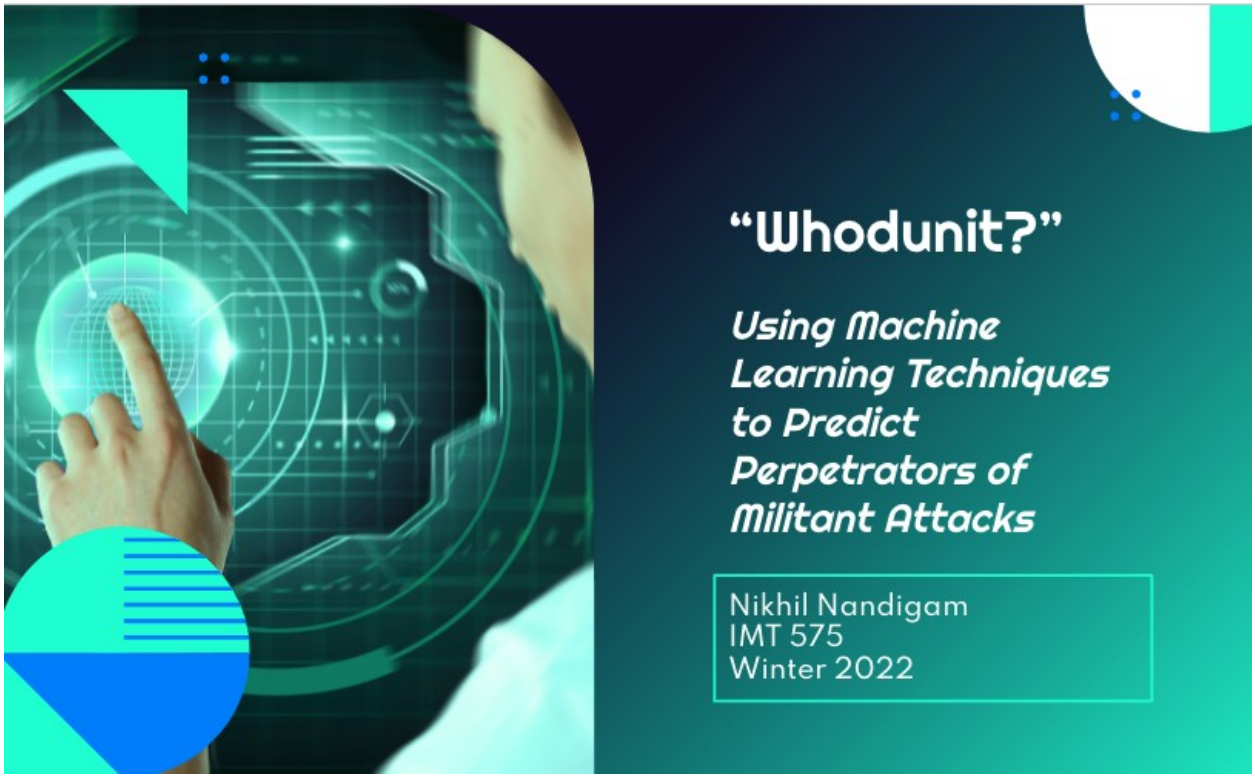


Disclaimer: This work does not advocate for operationalizing predictive techniques in the counter-terrorism domain. Rather, it serves as a thought experiment to complement rhetorical understandings of ISIS and al-Qaeda in the immediate aftermath of militant activity.

Question

Using the START-UMD Global Terrorism Database (GTD), can we configure a set of features to accurately predict whether an incident is conducted by an ISIS/ISIS-aligned group OR an al-Qaeda/al-Qaeda-aligned group?

Examining Non-Machine Learning Approaches

Studies into the differences between ISIS and al-Qaeda from fields of conflict and security studies focus on "doctrinal differences" based on public messaging by key ideologues from either camps (Arosoaie, 2015). Although there is much discourse around the motivations, rhetoric, and goals of these camps, this focus requires a great amount of domain expertise, linguistic skills, and historic context to access. Moreover, we do not gain an understanding as to how this rhetoric translates to on-the-ground incidents.

Non-machine learning approaches are then complemented by official attack claims for incidents. Although these claims perform a valuable role in confirming the perpetrators of incidents, groups release these claims up to 48-72 hours after an incident occurs, resulting in delayed response.

Existing Machine Learning Approaches

This work is not the first to apply machine learning techniques to the GTD observations. Xiaohui (2021) also conceptualizes a classification problem to exploit 36 features of the data to predict whether an incident is committed by one of 32 of the most represented militant groups in the dataset. Xiaohui conducts features engineering through a combination of domain expertise to exclude certain "subjective" features and an ExtraTrees classifier. Ultimately, this work attains a 97% accuracy using XGBoost and random forest classifiers.

Xiaohui's work serves as a starting point for my own classification problem. However, my model differs in that it serves a niche purpose to test whether the rhetorical distinction made between the ISIS and AQ camps can be reflected by the data. Furthermore, I examine the 36 features considered by Xiaohui. This allows for an operational model which predicts to which class an incident belongs while only using immediately ascertainable features, allowing users to deploy the model in the direct aftermath. Finally, while Xiaohui studies model predictions across a range of militant groups, I only focus on a subset of 47 groups, falling under either the ISIS or AQ ideological camps.

Key Results & Impact

High Perpetrator Predictability

At face value, we can predict with up to 95% accuracy whether a particular incident was carried out by a militant group ascribing to the ISIS camp or the AQ camp. However, a relatively high accuracy score across a complex problem space motivates us to challenge these results in the following sections.

Potential for Early Perpetrator Detection

By using features such as the country, region of the world, attack type, weapon type, and target type, our accurate predictions become more valuable given the dimension of making predictions which are attainable moments after an incident takes place, provided our model performs with similar accuracy scores across validation data.

Impact on Security and Conflict Studies

This question and analysis seeks to test whether the often-studied rhetorical dichotomy between ISIS ideology and AQ ideology translates into parsable tactical differences. Algorithmic visualizations suggest that two ideological camps are not ideal in clustering these two ideological camps. Perhaps ISIS is not always ISIS, but takes on a noticeable local flavor when in East Asia than when in the Sahel? Perhaps two South Asian militant groups, one under the AQ camp and the other under the ISIS camp, share more tactical similarities due to the local landscape? This discussion forms a quantitative starting point for further research into these questions.

Method

Dataset

The GTD contains various data from 1970-2019 pertaining to over 201,000 incidents of domestic and international terrorism, containing features such as the incident's location, attack type, relevant perpetrators, targets, and weapons involved. In total, the data contains 135 features. The incidents are collated through partnerships with multiple institutions, with citations for each incident gathered through a team of curators and machine learning tools.

The START-UMD data set includes incidents aligning with the following definition of terrorism: "The threatened or actual use of illegal force and violence by a non-state actor to attain a political, economic, religious, or social goal through fear, coercion, or intimidation (National)."

The data is currently curated by a team from the National Consortium for the Study of Terrorism and Responses to Terrorism and the University of Maryland.

Below we see five observations from the raw GTD database.

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.metrics import accuracy_score
from sklearn.metrics import confusion_matrix
from sklearn.model_selection import train_test_split
from sklearn.naive_bayes import MultinomialNB, CategoricalNB,
GaussianNB
from sklearn.svm import SVC
from sklearn.mixture import GaussianMixture
from sklearn.manifold import TSNE
import seaborn as sns

#Read GTD spreadsheet
gtd = pd.read_excel('/Users/nikhiln/Desktop/gtd.xlsx')

#Examine loaded entries
gtd.head()
```

```
        eventid  iyear  imonth  iday approxdate  extended resolution
country  \
0  197000000001   1970       7     2        NaN         0        NaT
58
1  197000000002   1970       0     0        NaN         0        NaT
130
2  197001000001   1970       1     0        NaN         0        NaT
160
3  197001000002   1970       1     0        NaN         0        NaT
78
4  197001000003   1970       1     0        NaN         0        NaT
101
```

```
          country_txt  region  ... addnotes scite1 scite2  scite3
dbsource  \
0  Dominican Republic       2  ...      NaN    NaN    NaN     NaN
PGIS
1              Mexico       1  ...      NaN    NaN    NaN     NaN
PGIS
2         Philippines       5  ...      NaN    NaN    NaN     NaN
PGIS
3              Greece       8  ...      NaN    NaN    NaN     NaN
PGIS
4               Japan       4  ...      NaN    NaN    NaN     NaN
PGIS

   INT_LOG  INT_IDEO INT_MISC INT_ANY  related
0        0         0        0       0      NaN
1        0         1        1       1      NaN
2       -9        -9        1       1      NaN
3       -9        -9        1       1      NaN
4       -9        -9        1       1      NaN

[5 rows x 135 columns]
```

Preprocessing

The first step in preprocessing is to filter the 200,000+ observations to only include those from January 2014 and onward. I apply this date given ISIS declared its self-proclaimed caliphate toward the end of June 2014 (“Timeline”). Its activities predate this declaration. However, setting a start date for analyzing incidents is helpful given al-Qaeda predates ISIS by a generation.

Our dataset now contains over 75,000 entries.

```
#Filter data for only entries from Jan 2014 onward
gtd_post_2014 = gtd[(gtd["iyear"] >= 2014) & (gtd["imonth"] >= 1)]

#Shape of data containing 75,831 incidents
gtd_post_2014.shape
```

```
(75831, 135)
```

The second step is to leverage domain expertise to review the 900+ unique group names contained in the post-2014 observations to create lists of groups falling into either the ISIS camp or AQ camp classes.

Below is an example of some of the militant group names occurring in the dataset before being categorized into one of the two camps. Groups which do not adhere to either of the two camps are excluded from the analysis.

For example, Somali militant group 'Al-Shabaab' is assigned to the AQ camp, while Sahel militant group 'Boko Haram' is assigned to the ISIS camp. Groups irrelevant to this project, such as 'Anarchists' or 'Buddhist Monks', are excluded.

```
#List unique values for perpetrator group names
group_names = gtd_post_2014.gname.unique()
print(len(group_names))

#Examples of group names
group_names[0:50]
```

```
903
```

```
array(['National Liberation Army of Colombia (ELN)', 'Al-Shabaab',
       'Jaish-e-Islam', 'Unknown', 'Muslim extremists',
       'Islamic State of Iraq and the Levant (ISIL)', 'Taliban',
       'Mozambique National Resistance Movement (MNR)', 'Maoists',
       'Lashkar-e-Balochistan',
       'Bangsamoro Islamic Freedom Movement (BIFM)', 'Separatists',
       'Al-Qaida in the Arabian Peninsula (AQAP)',
       'Allied Democratic Forces (ADF)',
       'Tehrik-i-Taliban Pakistan (TTP)', 'Jund al-Islam',
       'Ansar Bayt al-Maqdis (Ansar Jerusalem)', 'Seleka',
       'Baloch Republican Army (BRA)',
       'Communist Party of India - Maoist (CPI-Maoist)',
       'Abu Jaafar al-Mansur Brigades', 'Hadramawt Tribes Alliance',
       "Manipur Naga People's Army (MNPA)",
       'Bangladesh Nationalist Party (BNP)',
       'Houthi extremists (Ansar Allah)', 'Tribesmen',
       'Fulani extremists', 'Lashkar-e-Jhangvi', 'Anarchists',
       'Boko Haram', 'Abu Sayyaf Group (ASG)',
       'Achik National Volunteer Council-B (ANVC-B)',
       "New People's Army (NPA)", 'Punjabi Taliban',
       'Moro Islamic Liberation Front (MILF)', 'Caucasus Emirate',
       'Sudan Liberation Movement', 'Baloch Liberation Army (BLA)',
       'Revolutionary Armed Forces of Colombia (FARC)', 'Tabu Tribe',
       'Southern Mobility Movement (Yemen)',
       'Baloch Liberation Front (BLF)', "Lord's Resistance Army
(LRA)",
       'Democratic Front for the Liberation of Rwanda (FDLR)',
       'Achik Tiger Force', 'Sudan Liberation Army-Minni Minawi (SLA-
MM)',
       'Kuki National Front (KNF)', 'Abu Bakr Unis Jabr Brigade',
       'Group of Popular Fighters', 'Buddhist Monks'], dtype=object)
```

```
#Use domain expertise to categories group names into ISIS/pro-ISIS or
AQ/pro-AQ categories

ISIS = ['Ansar Bayt al-Maqdis (Ansar Jerusalem)',
        'Boko Haram',
        'Islamic State of Iraq and the Levant (ISIL)',
        'Abu Sayyaf Group (ASG)',
        'Bangsamoro Islamic Freedom Movement (BIFM)',
        'Mujahidin Indonesia Timur (MIT)',
```

```
        'Sinai Province of the Islamic State',
        'Barqa Province of the Islamic State',
        'Tripoli Province of the Islamic State',
        'Sanaa Province of the Islamic State',
        'Najd Province of the Islamic State',
        'Islamic State in Egypt',
        'Hijaz Province of the Islamic State',
        'Hadramawt Province of the Islamic State',
        'Islamic State in Bangladesh',
        'Algeria Province of the Islamic State'
        'Caucasus Province of the Islamic State',
        'Fezzan Province of the Islamic State',
        'Islamic State in the Greater Sahara (ISGS)',
        'Al Bayda Province of the Islamic State',
        'Jamaat Nusrat al-Islam wal Muslimin (JNIM)',
        'Ansar al-Sunna (Mozambique)',
        'East Asia Division of the Islamic State',
        'Central Africa Province of the Islamic State',
        'Mujahidin Indonesia Timur (MIT)',
        'Jamaat-ul-Ahrar',
        'Jundallah (Pakistan)']

AQ = ['Okba Ibn Nafaa Brigade',
      'Al-Qaida in the Indian Subcontinent',
      'Haqqani Network',
      "Hay'at Tahrir al-Sham",
      'Ansar Ghazwat-ul-Hind',
      'Ansar al-Tawhid',
      'Hurras al-Din',
      'Al-Shabaab',
      'Ansar al-Sharia (Libya)',
      'Al-Qaida in the Arabian Peninsula (AQAP)',
      'Al-Nusrah Front',
      'Lashkar-e-Taiba (LeT)',
      'Islamic Front (Syria)',
      'Liwa Ahrar al-Sunna',
      'Al-Qaida in the Islamic Maghreb (AQIM)',
      'Abdullah Azzam Brigades',
      "Al-Mua'qi'oon Biddam Brigade (Those who Sign with Blood)",
      'Turkestan Islamic Party',
      'Movement for Oneness and Jihad in West Africa (MUJAO)',
      'Ansarullah Bangla Team',
      'Ansar al-Islam (Egypt)']
```

After building lists to represent the individual groups which makeup each of my binary classes, I filter the post-2014 data again to only include incidents attributed to the groups of interest.

With our filtered GTD data set only containing incidents from 2014 and onward and carried out by groups falling under one of our two ideological camps, we can build a new column to represent our target. Our target will be represented as a binary "dummy" variable `is_isis`. If

our target is 0, the incident is attributed to a group under the AQ ideological camp. Otherwise, if the target value is 1, it falls under the ISIS camp.

```
#Create relevant_group consisting of 47 total groups falling within
two camps
relevant_groups = ISIS + AQ
print(len(relevant_groups))

#Filter data for only events occurring after January 2014 and
considered to fall inside either ISIS camp or AQ camp
gtd_filtered =
gtd_post_2014[gtd_post_2014['gname'].isin(relevant_groups)]

#Create dummy column for is_isis with 0 for AQ and 1 for ISIS
incidents
#https://stackify.dev/747481-python-generate-dummy-in-dataframe-based-
on-another-variable

gtd_filtered["is_isis"] = np.where(gtd_filtered["gname"].isin(ISIS),
1,0)
```

```
47

<ipython-input-5-c29b68e3d403>:11: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation:
https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#
returning-a-view-versus-a-copy
  gtd_filtered["is_isis"] = np.where(gtd_filtered["gname"].isin(ISIS),
1,0)
```

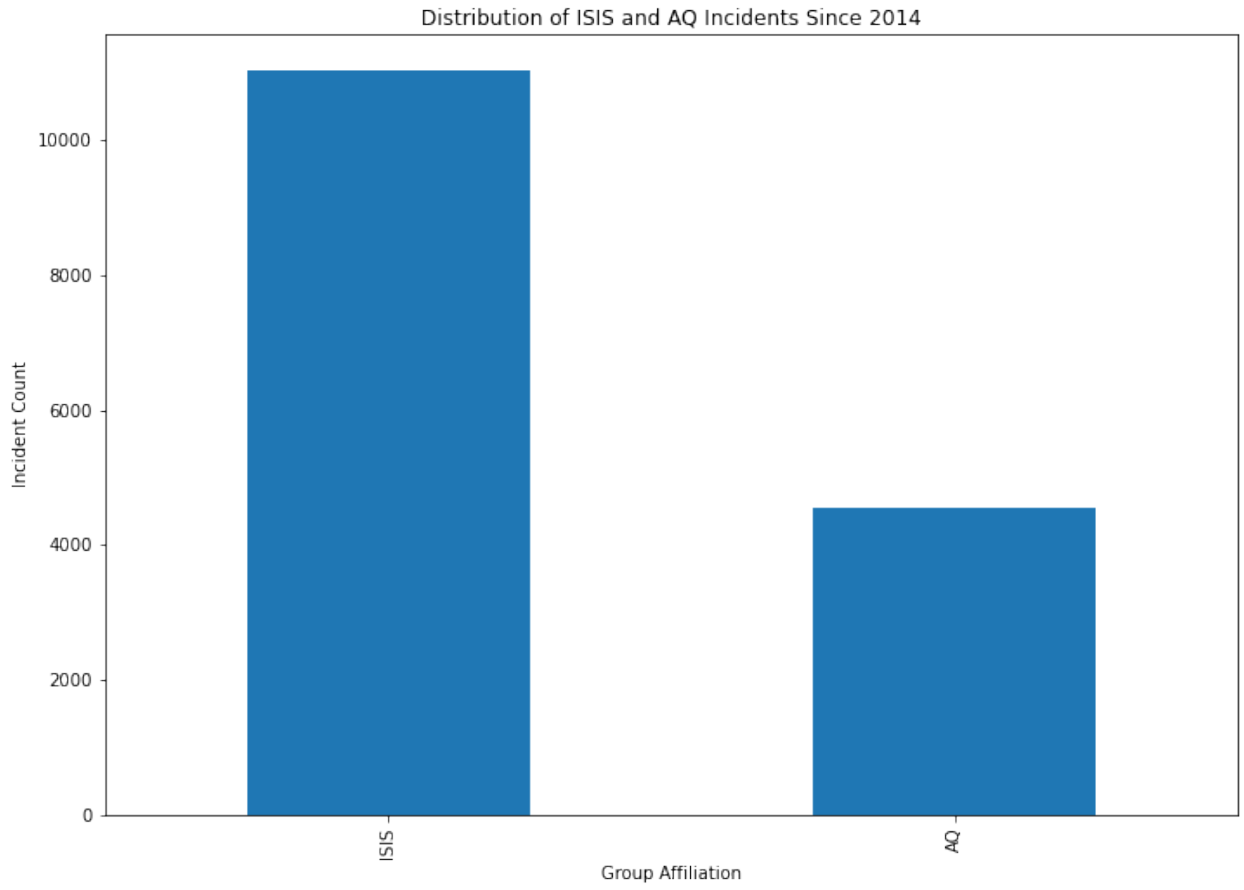
Before learning from our filtered data, it is worth examining the make up of our two classes. When considering the value counts of incidents in each class, we have an imbalanced classification limitation. Incidents attributed to the ISIS camp are represented at more than a 2:1 ratio compared to incidents belonging to our AQ camp class.

Below we see that our data contains 11,025 ISIS-affiliated incidents and 4,542 AQ-affiliated incidents between the years of 2014-2019. This class imbalance represents the on-the-ground reality of the nature of militant activity during this period. Rather than try to manipulate our filtered data to create a more balanced classification problem, we accept this reality as a limitation of the resulting model.

```
#Plot distributions of AQ and ISIS incidents
plt.figure(figsize=(11.7,8.27))

plot_counts = gtd_filtered['is_isis'].value_counts().plot(kind =
'bar')
```

```
x_labels = ["ISIS", "AQ"]

plot_counts.set_title("Distribution of ISIS and AQ Incidents Since
2014")
plot_counts.set_xlabel("Group Affiliation")
plot_counts.set_ylabel("Incident Count")
plot_counts.set_xticklabels(x_labels)


gtd_filtered['is_isis'].value_counts()
```

```
1    11025
0     4542
Name: is_isis, dtype: int64
```



Features & Models

As a first step toward building models, I chose features from the GTD which would be ascertainable in the immediate aftermath or even while an incident were taking place, allowing for the quickest understanding of the perpetrators of the event. Therefore, many of the 135 attributes available in the sparse GTD are irrelevant for my use case and are irrelevant to the type of incidents I am examining.

With this in mind, I used feature engineering by Xiaohui (2021) as a starting point to test the accuracy scores of predictions produced for the subset of militant groups within scope of this project.

Ultimately, my models contain five immediately ascertainable features: country, region of the world, primary target type, primary weapon type, and primary attack type.

Using a 70/30 train-test split, I built a categorical Naive Bayes model as my classification problem uses discrete features that are categorically distributed, with each category assigned to a unique number. The Naive Bayes classifier returns a 95% accuracy.

```
#Set features and target
X = gtd_filtered[['country', 'targtype1', 'weaptype1', 'region',
'attacktype1']]
y = gtd_filtered['is_isis']

#Use 70/30 train-test split
X_train, X_test, y_train, y_test = train_test_split(X, y,
test_size=0.3)

#Fit categorical Naive Bayes model
mnb = CategoricalNB()
mnb.fit(X_train,y_train.values.ravel())

#Make predictions
predictions_mnb = mnb.predict(X_test)

#Report accuracy and confusion matrix
print(accuracy_score(y_test,predictions_mnb))
print(confusion_matrix(y_test, predictions_mnb))
```

```
0.9554699207878399
[[1192  150]
 [  58 3271]]
```

I use a support vector machine model as my second predictive model. Although the categorical Naive Bayes model yielded a fairly high accuracy score across a complex problem space, this model assumes each feature is independent of the other. As we will discuss in the Limitations section, this is not the case with our features.

On the other hand, a polynomial kernel SVM model also considers the interactions between individual features to map them onto a non-linear space. This model performed less accurately than the Naive Bayes model at 87%, but this may be due to it learning more of the complexity of the problem than the Naive Bayes model.

```
#Support vector machine model
svm = SVC(kernel= 'poly')
svm.fit(X_train, y_train)

predictions_svm = svm.predict(X_test)
```

```
print(accuracy_score(y_test, predictions_svm))
print(confusion_matrix(y_test, predictions_svm))
```

```
0.867694283879255
[[ 928  414]
 [ 204 3125]]
```

Limitations

This analysis would be incomplete without complicating some of the high accuracy scores produced by the models.

Firstly, the class imbalance remains unresolved and it becomes easy for our models to overfit to the overrepresentation of ISIS-affiliated incidents in the training data. This imbalance also explains why both the Naive Bayes and SVM models perform better on predicting ISIS camp incidents over AQ camp incidents. The model is provided with more opportunities to learn from ISIS-attributed observations.

Secondly, another unresolved issue is the collinearity present in the predictors. An obvious instance of dependency is between the `country` and `region` features. A country will always be dependent on one and only one region of the world. There is also possible dependency between other predictors such as the `weaptype1` and `attacktype1`. For example, we may assume a strong correlation between explosives being used in explosions, or firearms used in assassinations.

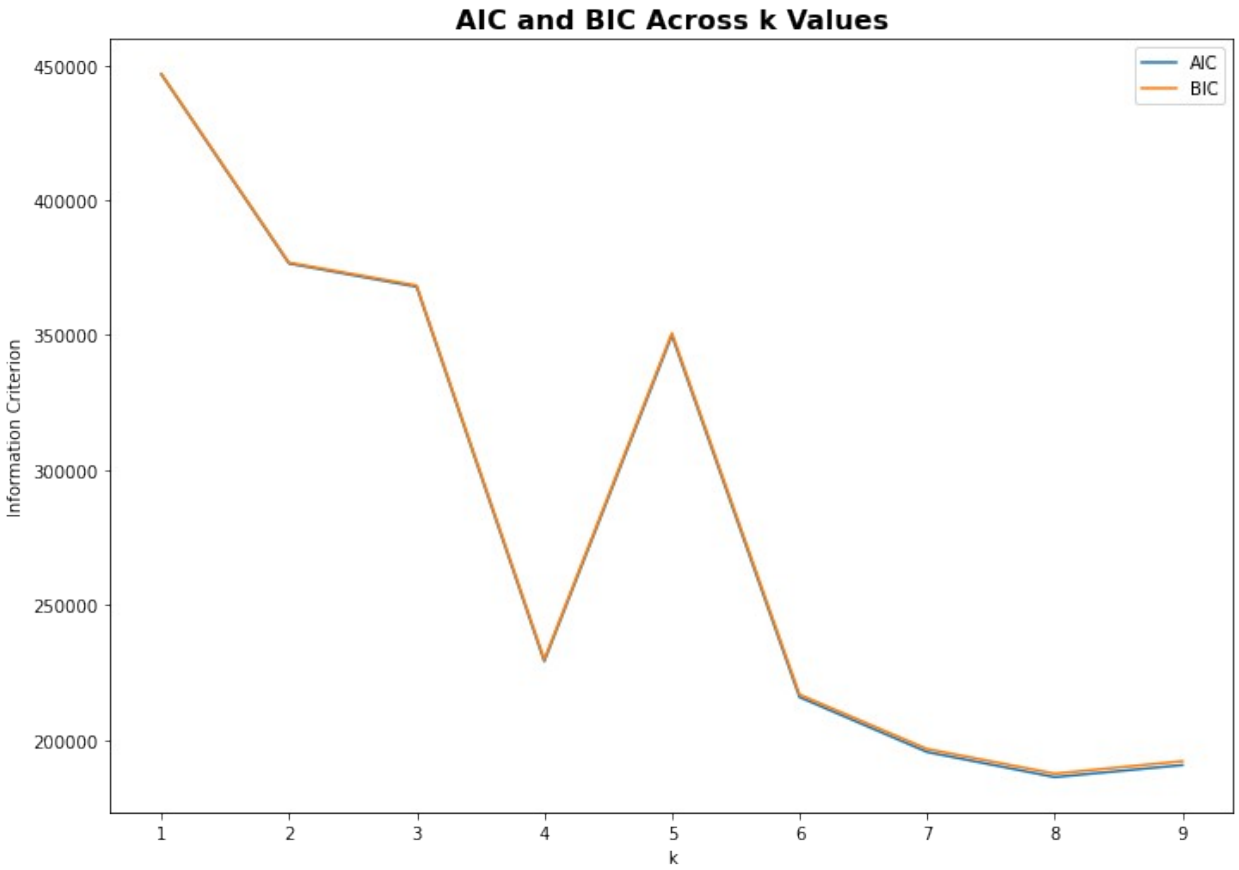
Shedding light on these limitations to our relatively high accuracy scores drives us to challenge one of the main assumptions of this analysis. Can we represent a clear distinction between the two ideological ISIS and AQ camps while considering only tangible attributes? If not, perhaps two classes are not the ideal number to model this problem.

To test this, I designed the visualization below to plot overlapping AIC and BIC values over a range of k (number of clusters) values to determine a better number of classes to model this problem. While seeking to minimize AIC and BIC values, the iteration of the model below suggests two classes do not ideally describe this problem space.

```
#Fit Gaussian Mixture EM model, considering two clusters (ISIS and AQ
camps)
model_EM = GaussianMixture(n_components=2, init_params='random',
max_iter=50)
model_EM.fit(X)

predictions_EM = model_EM.predict(X)

#Evaluate AIC and BIC values
print(model_EM.aic(X))
print(model_EM.bic(X))

#Examine resulting AIC and BIC values across range of other possible k
values
```

```
Sum_bic = []
Sum_aic = []

K = range(1, 10)

for k in K:
    gmm = GaussianMixture(n_components = k, init_params='random',
max_iter = 500)
    gmm = gmm.fit(X)
    Sum_bic.append(gmm.bic(X))
    Sum_aic.append(gmm.aic(X))

plt.figure(figsize=(11.7,8.27))
#Plot results
x1 = K
y1 = Sum_aic
plt.plot(x1, y1, label = "AIC")
x2 = K
y2 = Sum_bic
plt.plot(x2, y2, label = "BIC")


plt.title("AIC and BIC Across k Values", fontsize=16,
fontweight='bold')
plt.xlabel("k")
plt.ylabel("Information Criterion")
plt.legend(loc='upper right')
```

```
289955.1731972882
290268.94244857767

<matplotlib.legend.Legend at 0x7fa768f6aaf0>
```



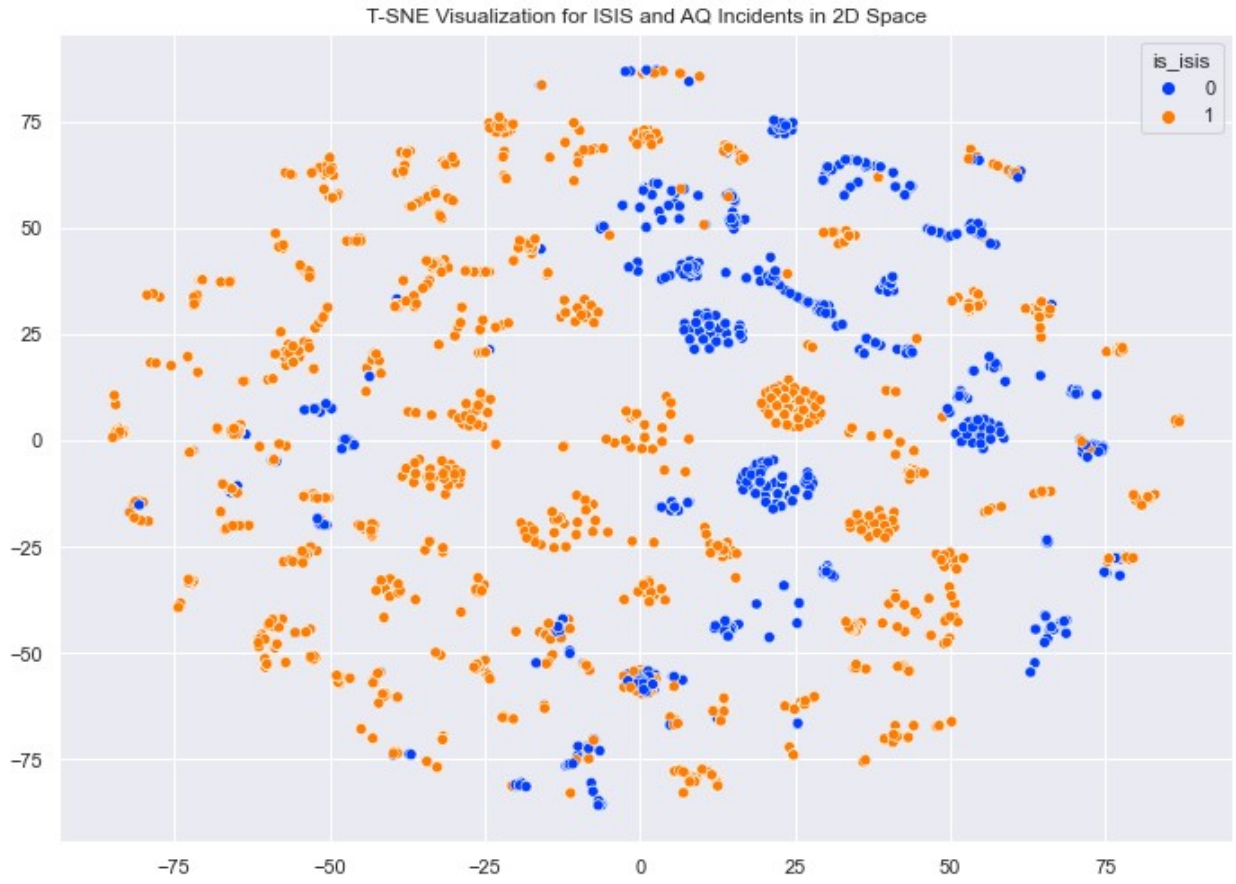
Below, I use a T-SNE plot to depict high-dimensional data over a two-dimensional space. The visualization further underscores the complexity of accurately predicting whether an incident falls into either the ISIS camp or the AQ camp.

The blue points represent incidents attributed to the AQ camp, and the orange represents ISIS camp incidents. Small clusters of each color lie fairly close to concentrations of the other color, and it becomes clear how a linear or even most polynomial models would not be able to easily classify this data.

```
#Fit high-dimension features to two-dimensional space using T-SNE
tsne = TSNE()
X_embedded = tsne.fit_transform(X)

#Plot incidents to investigate any trends/clusters
sns.set(rc={'figure.figsize':(11.7,8.27)})
palette = sns.color_palette("bright", 2)

sns.scatterplot(X_embedded[:,0], X_embedded[:,1], hue=y,
legend='full', palette=palette).set(title= 'T-SNE Visualization for
ISIS and AQ Incidents in 2D Space')
```

```
/Users/nikhiln/opt/anaconda3/lib/python3.8/site-packages/seaborn/
_decorators.py:36: FutureWarning: Pass the following variables as
keyword args: x, y. From version 0.12, the only valid positional
```

```
argument will be `data`, and passing other arguments without an
explicit keyword will result in an error or misinterpretation.
  warnings.warn(

[Text(0.5, 1.0, 'T-SNE Visualization for ISIS and AQ Incidents in 2D
Space')]
```



Risks & Ethical Considerations

As mentioned at the top of this discussion, the risks to operationalizing machine learning tools in the terrorism and militancy domain are numerous. Among them are the following concerns:

- Biased Data: The GTD set has been compiled to adhere to a single definition of terrorism. This definition results in the exclusion of potentially illegal activity conducted by state actors and militaries. The US government has sponsored some of the collection of the GTD set.

- Reliance on Data Entry and Citations: The GTD data is curated over decades by using human and machine learning efforts to capture citations of incidents worldwide. This results in a compounding effect of possible machine and human error or bias -- intentional or not. Naturally, incidents from media-dense environments will be reported. Other incidents may be unreported, misreported, or not collected by the

GTD curation team. Reporting of the details of incidents will reflect the source's own bias.

- Impact on Marginalized Communities: Different populations around the world have been disparately treated and disparately impacted by the 21st century's 'War on Terror'. There is a track record of dangers in military and intelligence sectors operationalizing machine learning against marginalized communities (Gibson, 2021). The impact on these communities should be prioritized and studied ahead of any legitimization of the predictions produced from these models.

Next Steps

This discussion serves as an opening to a larger effort in quantifying and visualizing tactical differences between militant groups which pledge loyalty to either the ISIS or AQ ideologies. Further steps include:

- Model Validation: Can we use details of militant incidents which occurred after 2019 to validate our model's predictions? This would not only test the predictive power of the models but also the robustness of the GTD curation team's efforts against data coded by a third party.
- Additional Feature Engineering: Can we model the differences between the two classes with a better set of features? Attaining data as provided by the GTD requires a costly human-managed process. The advantage is that we have features easily interpretable by humans and domain experts. However, could we use deep learning techniques to forgo this features engineering and produce better predictions at scale?
- Replicability in Other Conflict Contexts: Can any value be drawn from deploying these models to predict perpetrators of incidents in other conflict zones? Possible use cases include rivalries between cartel groups or other organized crime.

References

Arosoaie, A. (2015). Doctrinal Differences between ISIS and Al Qaeda: An Account of Ideologues. Counter Terrorist Trends and Analyses, 7(7), 31–37. http://www.jstor.org/stable/26351374

Gibson, J. (2021, February 18). Death by data: Drones, kill lists and algorithms. E-International Relations. Retrieved January 13, 2022, from https://www.e-ir.info/2021/02/18/death-by-data-drones-kill-lists-and-algorithms/

National Consortium for the Study of Terrorism and Responses to Terrorism. (n.d.). About the GTD. GTD. Retrieved January 13, 2022, from https://www.start.umd.edu/gtd/about/

Xiaohui Pan, "Quantitative Analysis and Prediction of Global Terrorist Attacks Based on Machine Learning", Scientific Programming, vol. 2021, Article ID 7890923, 15 pages, 2021. https://doi.org/10.1155/2021/7890923