From: Jeremy Heffner [jheffner@azavea.com]
Sent: Wednesday, September 16, 2015 11:56 AM
To: WILLIAMSON, DOUGLAS
Subject: Re: Accuracy evaluation project questions
Attachment(s): "2015_09_14 Azavea NYPD Agreement rev3.docx"

Doug --

We're fine with those changes. I made them within the attached document.

Jeremy

--
Jeremy Heffner

Azavea | 340 N 12th St, Suite 402, Philadelphia, PA
jheffner@azavea.com | T 215.701.7712 | F 215.925.2663
Web azavea.com | Blog azavea.com/blogs | Twitter @azavea

On Wed, Sep 16, 2015 at 10:46 AM, WILLIAMSON, DOUGLAS
<DOUGLAS.WILLIAMSON@nypd.org> wrote:

> Hi Jeremy.
>
> I think this is fine with the following exceptions. First, for comparative reasons, we
> cannot limit the test to one borough and will need to do the whole city. If this requires
> a larger grid than that is okay. And second, ideally we would get predictions per shift
> as this would provide the most utility operationally.
>
> Thanks.
>
> Doug
>
>
>
>
>
> **From:** Jeremy Heffner [mailto:jheffner@azavea.com]
> **Sent:** Monday, September 14, 2015 12:20 PM
> **To:** WILLIAMSON, DOUGLAS
>
>
> **Subject:** Re: Accuracy evaluation project questions
>
>
>
> Doug --
>
>
>
> Attached are revisions to the draft that you sent last week (tracked changes was on). I've tried
> to provide additional detail in areas where other evaluation processes we've been a part of
> became fuzzy or went awry.

Jeremy

--
Jeremy Heffner

Azavea | 340 N 12th St, Suite 402, Philadelphia, PA
jheffner@azavea.com | T 215.701.7712 | F 215.925.2663
Web azavea.com | Blog azavea.com/blogs | Twitter @azavea

On Fri, Sep 11, 2015 at 12:32 PM, WILLIAMSON, DOUGLAS
<DOUGLAS.WILLIAMSON@nypd.org> wrote:

Hi Jeremy,

Could they be delivered as GIS layers (ESRI feature classes)?

Doug

From: Jeremy Heffner [mailto:jheffner@azavea.com]
Sent: Friday, September 11, 2015 12:24 PM

To: WILLIAMSON, DOUGLAS
Subject: Re: Accuracy evaluation project questions

Doug --

At first glance, I am not seeing how the predictions will be conveyed to NYPD. Do you have a
sense of how that process will work?

Jeremy

--
Jeremy Heffner

Azavea | 340 N 12th St, Suite 402, Philadelphia, PA
jheffner@azavea.com | T 215.701.7712 | F 215.925.2663
Web azavea.com | Blog azavea.com/blogs | Twitter @azavea

On Fri, Sep 11, 2015 at 12:20 PM, WILLIAMSON, DOUGLAS
<DOUGLAS.WILLIAMSON@nypd.org> wrote:

Hi Jeremy,

Attached, please find the agreement for the NYPD - Azavea/HunchLab test project.

Please review, sign and return. Once signed, we will then sign and return to you.

Thanks!

Doug

Doug --

Attached is our data guide for HunchLab which may provide some ideas for you as well as information about what we can provide automatically. For purposes of a quickly accuracy demonstration I don't think we should try to have every data set that could possibly be relevant. Instead, I would focus on the ones that are readily available and that you believe correlate the most with criminal activity. As long as we have a nice, small basket of data points to represent different concepts things should work well.

For space-time event data, you may want to think about providing data broader than what you want us to model. For instance, lesser crime types that may be precursors, arrest data, etc. Any event that can be described as occurring at a specific XY coordinate at a specific time (or aoristic time range) can be used. Other data sets that may be interesting: 311 data, prison/probation releases, etc.

We can use any point, line, or polygon layer as geographic variables. We use the distance to the nearest feature and the density of features as variables within the models. ShapeFile format is the simplest to use. If there are attributes of the features that we should use to split the data set, we would just need this noted. For instance, if you have a parcel ShapeFile that includes zoning information as an attribute, we can split the layer into a layer per zoning type.

Temporal data would be in CSV format and could include data that has discrete states (on/off, severity levels) or continuous values. Examples may include: public school schedules, gang feuds (active feud between gang A and B, etc.), estimated attendance at special events, etc. The key part of temporal data is that we need to know historic values for the past 5 years in order to be able to use it within the model.

Jeremy

--
Jeremy Heffner

Azavea | 340 N 12th St, Suite 402, Philadelphia, PA
jheffner@azavea.com | T 215.701.7712 | F 215.925.2663
Web azavea.com | Blog azavea.com/blogs | Twitter @azavea

On Wed, Sep 9, 2015 at 2:42 PM, WILLIAMSON, DOUGLAS <DOUGLAS.WILLIAMSON@nypd.org> wrote:

Hi Jeremy,

No catalog. NYPD has TONs of GIS data but it isn't organized and it definitely isn't centralized. That is one of the things on my to-do list since I started here. That's why I was hoping you could provide us a list; that way I could just go through and if we have it, great, otherwise, maybe I could track it down.
Doug

From: Jeremy Heffner [mailto:jheffner@azavea.com]
Sent: Wednesday, September 09, 2015 2:20 PM
To: WILLIAMSON, DOUGLAS
Cc: Robert Cheetham

Subject: Re: Accuracy evaluation project questions

Doug --

Do you have a catalogue of data sets that we could browse to include in the agreement?

Robert will be signing the document as long as we receive it before he goes on vacation at the end of next week.

Robert Cheetham

President and CEO

Azavea

cheetham@azavea.com

215.701.7713

Jeremy

--
Jeremy Heffner

Azavea | 340 N 12th St, Suite 402, Philadelphia, PA
jheffner@azavea.com | T 215.701.7712 | F 215.925.2663
Web azavea.com | Blog azavea.com/blogs | Twitter @azavea

On Wed, Sep 9, 2015 at 1:15 PM, WILLIAMSON, DOUGLAS <DOUGLAS.WILLIAMSON@nypd.org> wrote:

Hi again.

Who will be signing the agreement, you or Robert Cheetham? If it is Robert, can I get his contact info?

Thanks,

Doug

From: Jeremy Heffner [mailto:jheffner@azavea.com]
Sent: Wednesday, September 09, 2015 12:04 PM
To: WILLIAMSON, DOUGLAS
Cc: John Branigan
Subject: Re: Accuracy evaluation project questions

Doug --

Is there some sort of data protection agreement that we need to execute? It would be good to get that underway. If we receive data next week, we will probably want at least two weeks for us to verify that things are working as expected before the evaluation period begins. We are doing team planning today after which I will have a better sense of our resources for the next few weeks.

In terms of data sets, I will get you some documentation about the data sets we would find particularly useful. We can also pull from open sources to fill the gaps.   I also still owe you some documentation about evaluation metrics and things we have learned from doing this

exercise previously.

Jeremy

--
Jeremy Heffner

Azavea | 340 N 12th St, Suite 402, Philadelphia, PA
jheffner@azavea.com | T 215.701.7712 | F 215.925.2663
Web azavea.com | Blog azavea.com/blogs | Twitter @azavea

On Wed, Sep 9, 2015 at 11:22 AM, WILLIAMSON, DOUGLAS
<DOUGLAS.WILLIAMSON@nypd.org> wrote:

Hi Jeremy,

We could probably get you the historic data by early next week. We do not have an
exact time frame other than 'as soon as possible'.

I assume you would need other data as well? Environmental data, etc? We have some
(liquor licenses, supermarkets, hospitals, etc). Some of it is fairly current, but some of
it may be somewhat out of date. Would all this be useful?

Doug

From: Jeremy Heffner [mailto:jheffner@azavea.com]
Sent: Thursday, September 03, 2015 3:01 PM
To: WILLIAMSON, DOUGLAS
Cc: John Branigan
Subject: Accuracy evaluation project questions

Hi Doug --

I've conferred with the team about the 30 days of predictions to determine our schedule for the
upcoming weeks.   We have a few questions:

When could the historic data be provided to us by? How much time from that point would we
have to provide predictions? We schedule work in 2 week blocks called sprints. The next
sprint begins on this coming Wednesday. We're trying to get a sense of when we would be
conducting the work to fit it into our existing commitments.

You mentioned that the predictions are across 30 days. Are those one step ahead forecasts or
forecasts across the entire period in one batch? For example, does that mean mean we have
data for 5 years ending on August 31st and then provide forecasts for September in one batch,

or do we receive updates for data in September as we submit each day's forecasts?

Jeremy

--
Jeremy Heffner

Azavea | 340 N 12th St, Suite 402, Philadelphia, PA
jheffner@azavea.com | T 215.701.7712 | F 215.925.2663
Web azavea.com | Blog azavea.com/blogs | Twitter @azavea

AZAVEA, INC.

PREDICTIVE ANALYTICS TEST PROJECT AGREEMENT

    The New York City Police Department, with headquarters at One Police Plaza, New York, New York 10038 ("NYPD") and Azavea, Inc., 340 N 12th St #402, Philadelphia, PA 19107 agree as follows:

1.   The NYPD is interested in conducting a test project of crime data predictive analytics.

2.   For this test project, the NYPD will provide five (5) years of crime complaint data, for the entirety of New York, consisting of the following information:

        a. Crime type;
        b. Latitude and Longitude;
        c. Date and Time of Occurrence;
        d. Precinct Location;
        e. Sector Location;
        f. Complaint Number.
        g. Last Updated date and time

      Dates and times will be provided in ISO 8601 format. Complaint Number values are unique record identifiers so that Azavea can de-duplicate records that have been updated in subsequent data extracts. Latitude and longitude information will reflect true coordinates and not be adjusted to mask the true location of events.

      f.   Data will be provided for all crime types so that precursor events can be used within predictive models.

3.   The crime complaint data will be provided to Azavea via a CSV file and transmitted via SFTP. Azavea will be responsible for extracting the data from the provided files and incorporating it into their analytics.

4.   In addition to the crime complaint data, additional crime correlate data will also be provided. These include, but are not limited to, the locations of:

        a.   Schools;
        b.   Hospitals;
        c.   Subway Entrances;
        d.   Mental Health Facilities;
        e.   Methadone Clinics;
        f.   Pawn Shops;
        g.   Liquor Licenses;
        h.   Restaurants;
        i.   Laundry Facilities;

Comment [JH1]: This makes managing data easier. If it is unavailable that is ok.

Formatted: Normal, No bullets or numbering

Comment [JH2]: Or replace with the relevant format.

Formatted: Normal, Indent: Before: 1", No bullets or numbering

j.   Parks.

5.   The location data will be provided to Azavea via an ESRI Geodatabase file and
     transmitted via SFTP. Azavea will be responsible for extracting the data from the
     provided database and incorporating it into their analytics. Azavea may utilize
     additional geographic and temporal data within the analysis including but not limited to
     weather, additional POI data sets, natural terrain data, etc. Azavea will bear any costs
     related to these other data sets.

6.   ~~Azavea will perform its analytical function using the NYPD data and provide predictions
     each day to the NYPD.~~ The NYPD will provide a CSV file each day by 9AM via the same
     SFTP transmission method with any changes and new entries that occurred since the
     last report, using the same procedure described in section 3. This data will be
     incorporated into ~~its~~ the analysis and used to provide the next set of predictions. Azavea
     will perform its analytical function using the NYPD data and provide predictions each
     day to the NYPD by noon. Predictions will be provided to the NYPD via email in
     GeoJSON or ShapeFile format.

7.   Azavea's predictions will cover the highest risk 1% of the land area for the geographic
     area being modeled (the city). Predictions will be made for each shift of the day. Azavea
     will provide two sets of predictions. Set one will combine predictions for different crime
     types into one set of target areas based upon a NYPD determined weighting scheme
     between crime types. Set two will provide separate prediction layers for each crime
     type so that accuracy can be measured for different types of events. Each layer of
     predictions within set two will cover 1% of the land area.

8~~7~~.   The predictive analytics test project is at no cost to the NYPD.

~~8.   Azavea will conduct the predictive analytics test for no more than thirty (30) days.~~

9.   At the end of the thirty (30) day testing period plus a six-month grace period to provide
     for mutual analysis of the results, Azavea must permanently delete or destroy all copies
     of NYPD data, predictions, and analytical reports. The data must not be used for any
     purpose other than providing predictive analysis to the NYPD during the term of the
     project. The predictions and information contained in any analytical reports must not
     be used by Azavea for any reason other than this test project.

10. To protect the accuracy of the predictions, the NYPD may not use the provided
     predictions for any operational decision making. Any evaluation metrics used to
     compare the accuracy of the Azavea predictions with other systems (internal or

external) will be designed to provide confidence intervals (such as from bootstrap sampling or tests of statistically significant differences in results) at the 90% level or greater. Evaluation metrics for the combined crime type predictions will utilize the same weighting scheme that NYPD provided to Azavea. Evaluation metrics for Azavea's predictions and all (anonymized) systems the predictions are compared against will be provided to Azavea for review. Evaluation metrics will be kept confidential by the NYPD unless granted written permission to release the metrics by Azavea. Azavea will also keep evaluation metrics confidential unless granted written permission to release the metrics by NYPD.

110.   Azavea acknowledges that the NYPD has made no promise, express or implied, concerning the possible future purchase of predictive analytics services and that the NYPD is under no obligation to purchase predictive analytics services as a result of the agreement. All NYPD procurements are conducted in accordance with the New York City Procurement Policy Board Rules.

121.   Azavea and the NYPD designate these individuals to serve as the points of contact for issues related to this Agreement:

Azavea: Jeremy Heffner
            Senior Data Scientist & Product Manager
            Office: 215.701.7712
            Email: jheffner@azavea.com

NYPD: Douglas Williamson
            Director, Operations Research
            Management Analysis and Planning
            Office: 646-610-5076
            Email: douglas.williamson@nypd.org

132.   This Agreement represents the entire agreement between Azavea, Inc. and the NYPD concerning the provision of predictive analytics as a test project. Any amendments to this agreement shall be in writing, signed by both parties.

In witness thereof, the parties have executed this Agreement by the signatures of duly authorized officials on the _____ day of September, 2015.

AZAVEA, Inc.

NEW YORK CITY
POLICE DEPARTMENT

_____   _____

Robert Cheetham   Ronald Wilhelmy
President and CEO   Assistant Commissioner
AZAVEA, Inc.   NYPD, Management Analysis and Planning

From: Jeremy Heffner [jheffner@azavea.com]
Sent: Friday, September 18, 2015 2:11 PM
To: WILLIAMSON, DOUGLAS
Subject: Re: Accuracy evaluation project questions

One other point I failed to make:

You can use this type of analysis to get a sense of how much of a difference would be detectable based upon your typical crime levels.

For example, if over the course of a month you typically have 1000 burglaries and you assume that a good model would capture 15% in the top 1% of the land area, you could proactively enter that into the calculation and see that the confidence intervals would be 13.27 to 16.97. A second model that captures 10% would lead to intervals of 8.58 to 11.67. In such a case you could say that one is better than the other.

If instead, you only had 100 burglaries over a month, then models that caught 15% and 10% of the events would lead to confidence intervals of 5.88 to 15.68 and 9.8 to 21.57 respectively. In that case you wouldn't be able to say that one model is better than the other.

Jeremy

--
Jeremy Heffner

Azavea | 340 N 12th St, Suite 402, Philadelphia, PA
jheffner@azavea.com | T 215.701.7712 | F 215.925.2663
Web azavea.com | Blog azavea.com/blogs | Twitter @azavea

On Fri, Sep 18, 2015 at 3:02 PM, Jeremy Heffner <jheffner@azavea.com> wrote:
> Hey Doug --
>
> I finally got time to comment up the code I have to give you an example.
>
> There are often significance tests that can be applied to metrics (like detecting differences in means), but I've leaned towards bootstrap estimation because it provides a valid approach to estimating confidence intervals even for complicated metrics or where you are unsure that data fits a certain distribution, etc.
>
> The Predictive Accuracy Index (PAI) is one metric put forward by Spencer Chainey to compare the predictive power of different geographic recommendation systems. The metric is simply (percent events captured) / (percent land flagged). If the percent of land area flagged is fixed between the systems being compared, then this metric reduces to the percent of events captured.
>
> Note that it is possible for vendors to game this metric by optimizing the percent area flagged to maximize the metric. It is best to have different approaches identify the same percent area. The metric is also difficult to compare across different data sets because it does not account for the natural concentration of a given data set. Nonetheless it's a commonly used metric.
>
> The attached script shows an example of how to form confidence intervals for PAI based upon bootstrap sampling. For example, let's assume we capture 5 out of 10 events in 1% of the land area. This leads to a PAI of 50. Each bootstrap sample will sample (with replacement) from the 10 events where 5 are marked as captured (TRUE in the script) -- some samples will show more events captured, some less. The PAI is calculated on each sample and then the center

90% of the sampled values forms the confidence interval.

The script also adjusts the captured and total counts with a balanced set of two hypothetical events -- one captured, one not captured. This enables calculation of confidence intervals even in cases where the actual values captured none or all events.   Essentially this slightly widens the confidence intervals by including the next event that may or may not have been captured (the two extremes).   As the real counts increase these extra events change the estimate less and less.

Jeremy

--
Jeremy Heffner

Azavea | 340 N 12th St, Suite 402, Philadelphia, PA
jheffner@azavea.com | T 215.701.7712 | F 215.925.2663
Web azavea.com | Blog azavea.com/blogs | Twitter @azavea

On Wed, Sep 16, 2015 at 10:49 AM, WILLIAMSON, DOUGLAS <DOUGLAS.WILLIAMSON@nypd.org> wrote:

> Hi again.
>
> Would you be able to provide some resources on the evaluation metrics? I have come across some basic ones, but none related to the bootstrapping methods you reference.
>
> Thanks again,
>
> Doug
>
>
>
> **From:** Jeremy Heffner [mailto:jheffner@azavea.com]
> **Sent:** Monday, September 14, 2015 12:20 PM
> **To:** WILLIAMSON, DOUGLAS
>
> **Subject:** Re: Accuracy evaluation project questions
>
>
> Doug --
>
>
> Attached are revisions to the draft that you sent last week (tracked changes was on). I've tried to provide additional detail in areas where other evaluation processes we've been a part of became fuzzy or went awry.
>
>
> Jeremy
>
> --

Jeremy Heffner

Azavea | 340 N 12th St, Suite 402, Philadelphia, PA
jheffner@azavea.com | T 215.701.7712 | F 215.925.2663
Web azavea.com | Blog azavea.com/blogs | Twitter @azavea

On Fri, Sep 11, 2015 at 12:32 PM, WILLIAMSON, DOUGLAS
<DOUGLAS.WILLIAMSON@nypd.org> wrote:

Hi Jeremy,

Could they be delivered as GIS layers (ESRI feature classes)?

Doug

From: Jeremy Heffner [mailto:jheffner@azavea.com]
Sent: Friday, September 11, 2015 12:24 PM

To: WILLIAMSON, DOUGLAS
Subject: Re: Accuracy evaluation project questions

Doug --

At first glance, I am not seeing how the predictions will be conveyed to NYPD. Do you have a sense of how that process will work?

Jeremy

--
Jeremy Heffner

Azavea | 340 N 12th St, Suite 402, Philadelphia, PA
jheffner@azavea.com | T 215.701.7712 | F 215.925.2663
Web azavea.com | Blog azavea.com/blogs | Twitter @azavea

On Fri, Sep 11, 2015 at 12:20 PM, WILLIAMSON, DOUGLAS
<DOUGLAS.WILLIAMSON@nypd.org> wrote:

Hi Jeremy,

Attached, please find the agreement for the NYPD - Azavea/HunchLab test project.

Please review, sign and return. Once signed, we will then sign and return to you.

Thanks!

Doug

Subject: Re: Accuracy evaluation project questions

Doug --

Attached is our data guide for HunchLab which may provide some ideas for you as well as information about what we can provide automatically. For purposes of a quickly accuracy demonstration I don't think we should try to have every data set that could possibly be relevant. Instead, I would focus on the ones that are readily available and that you believe correlate the most with criminal activity. As long as we have a nice, small basket of data points to represent different concepts things should work well.

For space-time event data, you may want to think about providing data broader than what you want us to model. For instance, lesser crime types that may be precursors, arrest data, etc. Any event that can be described as occurring at a specific XY coordinate at a specific time (or aoristic time range) can be used. Other data sets that may be interesting: 311 data, prison/probation releases, etc.

We can use any point, line, or polygon layer as geographic variables. We use the distance to the nearest feature and the density of features as variables within the models. ShapeFile format is the simplest to use. If there are attributes of the features that we should use to split the data set, we would just need this noted. For instance, if you have a parcel ShapeFile that includes zoning information as an attribute, we can split the layer into a layer per zoning type.

Temporal data would be in CSV format and could include data that has discrete states (on/off, severity levels) or continuous values. Examples may include: public school schedules, gang feuds (active feud between gang A and B, etc.), estimated attendance at special events, etc. The key part of temporal data is that we need to know historic values for the past 5 years in order to be able to use it within the model.

Jeremy

--
Jeremy Heffner

Azavea | 340 N 12th St, Suite 402, Philadelphia, PA
jheffner@azavea.com | T 215.701.7712 | F 215.925.2663
Web azavea.com | Blog azavea.com/blogs | Twitter @azavea

On Wed, Sep 9, 2015 at 2:42 PM, WILLIAMSON, DOUGLAS
<DOUGLAS.WILLIAMSON@nypd.org> wrote:

Hi Jeremy,

No catalog. NYPD has TONs of GIS data but it isn't organized and it definitely isn't
centralized. That is one of the things on my to-do list since I started here. That's why
I was hoping you could provide us a list; that way I could just go through and if we
have it, great, otherwise, maybe I could track it down.
Doug

From: Jeremy Heffner [mailto:jheffner@azavea.com]
Sent: Wednesday, September 09, 2015 2:20 PM
To: WILLIAMSON, DOUGLAS
Cc: Robert Cheetham

Subject: Re: Accuracy evaluation project questions

Doug --

Do you have a catalogue of data sets that we could browse to include in the agreement?

Robert will be signing the document as long as we receive it before he goes on vacation at
the end of next week.

Robert Cheetham

President and CEO

Azavea

cheetham@azavea.com

Jeremy

--
Jeremy Heffner

Azavea | 340 N 12th St, Suite 402, Philadelphia, PA
jheffner@azavea.com | T 215.701.7712 | F 215.925.2663
Web azavea.com | Blog azavea.com/blogs | Twitter @azavea

On Wed, Sep 9, 2015 at 1:15 PM, WILLIAMSON, DOUGLAS <DOUGLAS.WILLIAMSON@nypd.org> wrote:

Hi again.

Who will be signing the agreement, you or Robert Cheetham? If it is Robert, can I get his contact info?

Thanks,

Doug

From: Jeremy Heffner [mailto:jheffner@azavea.com]
Sent: Wednesday, September 09, 2015 12:04 PM
To: WILLIAMSON, DOUGLAS
Cc: John Branigan
Subject: Re: Accuracy evaluation project questions

Doug --

Is there some sort of data protection agreement that we need to execute? It would be good to get that underway. If we receive data next week, we will probably want at least two weeks for us to verify that things are working as expected before the evaluation period begins. We are doing team planning today after which I will have a better sense of our resources for the next few weeks.

In terms of data sets, I will get you some documentation about the data sets we would find particularly useful. We can also pull from open sources to fill the gaps.   I also still owe you some documentation about evaluation metrics and things we have learned from doing this exercise previously.

Jeremy

--
Jeremy Heffner

Azavea | 340 N 12th St, Suite 402, Philadelphia, PA
jheffner@azavea.com | T 215.701.7712 | F 215.925.2663
Web azavea.com | Blog azavea.com/blogs | Twitter @azavea

On Wed, Sep 9, 2015 at 11:22 AM, WILLIAMSON, DOUGLAS
<DOUGLAS.WILLIAMSON@nypd.org> wrote:

Hi Jeremy,

We could probably get you the historic data by early next week. We do not have an
exact time frame other than 'as soon as possible'.

I assume you would need other data as well? Environmental data, etc? We have
some (liquor licenses, supermarkets, hospitals, etc). Some of it is fairly current, but
some of it may be somewhat out of date. Would all this be useful?

Doug

From: Jeremy Heffner [mailto:jheffner@azavea.com]
Sent: Thursday, September 03, 2015 3:01 PM
To: WILLIAMSON, DOUGLAS
Cc: John Branigan
Subject: Accuracy evaluation project questions

Hi Doug --

I've conferred with the team about the 30 days of predictions to determine our schedule for
the upcoming weeks.   We have a few questions:

When could the historic data be provided to us by? How much time from that point would we
have to provide predictions? We schedule work in 2 week blocks called sprints. The next
sprint begins on this coming Wednesday. We're trying to get a sense of when we would be
conducting the work to fit it into our existing commitments.

You mentioned that the predictions are across 30 days. Are those one step ahead forecasts
or forecasts across the entire period in one batch? For example, does that mean mean we
have data for 5 years ending on August 31st and then provide forecasts for September in
one batch, or do we receive updates for data in September as we submit each day's
forecasts?

Jeremy

--
Jeremy Heffner

Azavea | 340 N 12th St, Suite 402, Philadelphia, PA
jheffner@azavea.com | T 215.701.7712 | F 215.925.2663
Web azavea.com | Blog azavea.com/blogs | Twitter @azavea

From: Jeremy Heffner [jheffner@azavea.com]
Sent: Friday, September 18, 2015 2:11 PM
To: WILLIAMSON, DOUGLAS
Subject: Re: Accuracy evaluation project questions

One other point I failed to make:

You can use this type of analysis to get a sense of how much of a difference would be detectable based upon your typical crime levels.

For example, if over the course of a month you typically have 1000 burglaries and you assume that a good model would capture 15% in the top 1% of the land area, you could proactively enter that into the calculation and see that the confidence intervals would be 13.27 to 16.97.   A second model that captures 10% would lead to intervals of 8.58 to 11.67.   In such a case you could say that one is better than the other.

If instead, you only had 100 burglaries over a month, then models that caught 15% and 10% of the events would lead to confidence intervals of 5.88 to 15.68 and 9.8 to 21.57 respectively. In that case you wouldn't be able to say that one model is better than the other.

Jeremy

--
Jeremy Heffner

Azavea | 340 N 12th St, Suite 402, Philadelphia, PA
jheffner@azavea.com | T 215.701.7712 | F 215.925.2663
Web azavea.com | Blog azavea.com/blogs | Twitter @azavea

On Fri, Sep 18, 2015 at 3:02 PM, Jeremy Heffner <jheffner@azavea.com> wrote:
> Hey Doug --
>
> I finally got time to comment up the code I have to give you an example.
>
> There are often significance tests that can be applied to metrics (like detecting differences in means), but I've leaned towards bootstrap estimation because it provides a valid approach to estimating confidence intervals even for complicated metrics or where you are unsure that data fits a certain distribution, etc.
>
> The Predictive Accuracy Index (PAI) is one metric put forward by Spencer Chainey to compare the predictive power of different geographic recommendation systems. The metric is simply (percent events captured) / (percent land flagged). If the percent of land area flagged is fixed between the systems being compared, then this metric reduces to the percent of events captured.
>
> Note that it is possible for vendors to game this metric by optimizing the percent area flagged to maximize the metric. It is best to have different approaches identify the same percent area. The metric is also difficult to compare across different data sets because it does not account for the natural concentration of a given data set. Nonetheless it's a commonly used metric.
>
> The attached script shows an example of how to form confidence intervals for PAI based upon bootstrap sampling. For example, let's assume we capture 5 out of 10 events in 1% of the land area. This leads to a PAI of 50.   Each bootstrap sample will sample (with replacement) from the 10 events where 5 are marked as captured (TRUE in the script) -- some samples will show more events captured, some less. The PAI is calculated on each sample and then the center

90% of the sampled values forms the confidence interval.

The script also adjusts the captured and total counts with a balanced set of two hypothetical events -- one captured, one not captured. This enables calculation of confidence intervals even in cases where the actual values captured none or all events.   Essentially this slightly widens the confidence intervals by including the next event that may or may not have been captured (the two extremes).   As the real counts increase these extra events change the estimate less and less.

Jeremy

--
Jeremy Heffner

Azavea | 340 N 12th St, Suite 402, Philadelphia, PA
jheffner@azavea.com | T 215.701.7712 | F 215.925.2663
Web azavea.com | Blog azavea.com/blogs | Twitter @azavea

On Wed, Sep 16, 2015 at 10:49 AM, WILLIAMSON, DOUGLAS <DOUGLAS.WILLIAMSON@nypd.org> wrote:

> Hi again.
>
> Would you be able to provide some resources on the evaluation metrics? I have come across some basic ones, but none related to the bootstrapping methods you reference.
>
> Thanks again,
>
> Doug
>
>
>
> **From:** Jeremy Heffner [mailto:jheffner@azavea.com]
> **Sent:** Monday, September 14, 2015 12:20 PM
> **To:** WILLIAMSON, DOUGLAS
>
>
> **Subject:** Re: Accuracy evaluation project questions
>
>
> Doug --
>
>
> Attached are revisions to the draft that you sent last week (tracked changes was on). I've tried to provide additional detail in areas where other evaluation processes we've been a part of became fuzzy or went awry.
>
>
> Jeremy
>
> --

Jeremy Heffner

Azavea | 340 N 12th St, Suite 402, Philadelphia, PA
jheffner@azavea.com | T 215.701.7712 | F 215.925.2663
Web azavea.com | Blog azavea.com/blogs | Twitter @azavea

On Fri, Sep 11, 2015 at 12:32 PM, WILLIAMSON, DOUGLAS
<DOUGLAS.WILLIAMSON@nypd.org> wrote:

Hi Jeremy,

Could they be delivered as GIS layers (ESRI feature classes)?

Doug

From: Jeremy Heffner [mailto:jheffner@azavea.com]
Sent: Friday, September 11, 2015 12:24 PM

To: WILLIAMSON, DOUGLAS
Subject: Re: Accuracy evaluation project questions

Doug --

At first glance, I am not seeing how the predictions will be conveyed to NYPD. Do you have a sense of how that process will work?

Jeremy

--
Jeremy Heffner

Azavea | 340 N 12th St, Suite 402, Philadelphia, PA
jheffner@azavea.com | T 215.701.7712 | F 215.925.2663
Web azavea.com | Blog azavea.com/blogs | Twitter @azavea

On Fri, Sep 11, 2015 at 12:20 PM, WILLIAMSON, DOUGLAS
<DOUGLAS.WILLIAMSON@nypd.org> wrote:

Hi Jeremy,

Attached, please find the agreement for the NYPD - Azavea/HunchLab test project.

Please review, sign and return. Once signed, we will then sign and return to you.

Thanks!

Doug

Subject: Re: Accuracy evaluation project questions

Doug --

Attached is our data guide for HunchLab which may provide some ideas for you as well as information about what we can provide automatically. For purposes of a quickly accuracy demonstration I don't think we should try to have every data set that could possibly be relevant. Instead, I would focus on the ones that are readily available and that you believe correlate the most with criminal activity. As long as we have a nice, small basket of data points to represent different concepts things should work well.

For space-time event data, you may want to think about providing data broader than what you want us to model. For instance, lesser crime types that may be precursors, arrest data, etc. Any event that can be described as occurring at a specific XY coordinate at a specific time (or aoristic time range) can be used. Other data sets that may be interesting: 311 data, prison/probation releases, etc.

We can use any point, line, or polygon layer as geographic variables. We use the distance to the nearest feature and the density of features as variables within the models. ShapeFile format is the simplest to use. If there are attributes of the features that we should use to split the data set, we would just need this noted. For instance, if you have a parcel ShapeFile that includes zoning information as an attribute, we can split the layer into a layer per zoning type.

Temporal data would be in CSV format and could include data that has discrete states (on/off, severity levels) or continuous values. Examples may include: public school schedules, gang feuds (active feud between gang A and B, etc.), estimated attendance at special events, etc. The key part of temporal data is that we need to know historic values for the past 5 years in order to be able to use it within the model.

Jeremy

--
Jeremy Heffner

Azavea | 340 N 12th St, Suite 402, Philadelphia, PA
jheffner@azavea.com | T 215.701.7712 | F 215.925.2663
Web azavea.com | Blog azavea.com/blogs | Twitter @azavea

On Wed, Sep 9, 2015 at 2:42 PM, WILLIAMSON, DOUGLAS
<DOUGLAS.WILLIAMSON@nypd.org> wrote:

Hi Jeremy,

No catalog. NYPD has TONs of GIS data but it isn't organized and it definitely isn't
centralized. That is one of the things on my to-do list since I started here. That's why
I was hoping you could provide us a list; that way I could just go through and if we
have it, great, otherwise, maybe I could track it down.
Doug

From: Jeremy Heffner [mailto:jheffner@azavea.com]
Sent: Wednesday, September 09, 2015 2:20 PM
To: WILLIAMSON, DOUGLAS
Cc: Robert Cheetham

Subject: Re: Accuracy evaluation project questions

Doug --

Do you have a catalogue of data sets that we could browse to include in the agreement?

Robert will be signing the document as long as we receive it before he goes on vacation at
the end of next week.

Robert Cheetham

President and CEO

Azavea

cheetham@azavea.com

Jeremy

--
Jeremy Heffner

Azavea | 340 N 12th St, Suite 402, Philadelphia, PA
jheffner@azavea.com | T 215.701.7712 | F 215.925.2663
Web azavea.com | Blog azavea.com/blogs | Twitter @azavea

On Wed, Sep 9, 2015 at 1:15 PM, WILLIAMSON, DOUGLAS <DOUGLAS.WILLIAMSON@nypd.org> wrote:

Hi again.

Who will be signing the agreement, you or Robert Cheetham? If it is Robert, can I get his contact info?

Thanks,

Doug

From: Jeremy Heffner [mailto:jheffner@azavea.com]
Sent: Wednesday, September 09, 2015 12:04 PM
To: WILLIAMSON, DOUGLAS
Cc: John Branigan
Subject: Re: Accuracy evaluation project questions

Doug --

Is there some sort of data protection agreement that we need to execute? It would be good to get that underway. If we receive data next week, we will probably want at least two weeks for us to verify that things are working as expected before the evaluation period begins. We are doing team planning today after which I will have a better sense of our resources for the next few weeks.

In terms of data sets, I will get you some documentation about the data sets we would find particularly useful. We can also pull from open sources to fill the gaps.   I also still owe you some documentation about evaluation metrics and things we have learned from doing this exercise previously.

Jeremy

--
Jeremy Heffner

Azavea | 340 N 12th St, Suite 402, Philadelphia, PA
jheffner@azavea.com | T 215.701.7712 | F 215.925.2663
Web azavea.com | Blog azavea.com/blogs | Twitter @azavea

On Wed, Sep 9, 2015 at 11:22 AM, WILLIAMSON, DOUGLAS
<DOUGLAS.WILLIAMSON@nypd.org> wrote:

Hi Jeremy,

We could probably get you the historic data by early next week. We do not have an
exact time frame other than 'as soon as possible'.

I assume you would need other data as well? Environmental data, etc? We have
some (liquor licenses, supermarkets, hospitals, etc). Some of it is fairly current, but
some of it may be somewhat out of date. Would all this be useful?

Doug

From: Jeremy Heffner [mailto:jheffner@azavea.com]
Sent: Thursday, September 03, 2015 3:01 PM
To: WILLIAMSON, DOUGLAS
Cc: John Branigan
Subject: Accuracy evaluation project questions

Hi Doug --

I've conferred with the team about the 30 days of predictions to determine our schedule for
the upcoming weeks.   We have a few questions:

When could the historic data be provided to us by? How much time from that point would we
have to provide predictions? We schedule work in 2 week blocks called sprints. The next
sprint begins on this coming Wednesday. We're trying to get a sense of when we would be
conducting the work to fit it into our existing commitments.

You mentioned that the predictions are across 30 days. Are those one step ahead forecasts
or forecasts across the entire period in one batch? For example, does that mean mean we
have data for 5 years ending on August 31st and then provide forecasts for September in
one batch, or do we receive updates for data in September as we submit each day's
forecasts?

Jeremy

--
Jeremy Heffner

Azavea | 340 N 12th St, Suite 402, Philadelphia, PA
jheffner@azavea.com | T 215.701.7712 | F 215.925.2663
Web azavea.com | Blog azavea.com/blogs | Twitter @azavea