$\underline{(https://programs.upgrad.com/data-science-advanced-certificate-bdm-iimk/?}\\$

utm_source=BRAND&utm_medium=AIM&utm_campaign=DV_DA_IIMK_BRAND_AIM_MARCH5-ROADBLOCK_METRO_WEBSITE)



Innovative Platform to learn Data Science
150,000+ Data Enthusiasts Globally



 $\underline{(https://leaps.analyttica.com/innovative-datascience-learning-platform?}$

utm_source=analyticsindiamagazine&utm_medium=partner_website&utm_campaign=aim_learndatascience_partnership)

DEVELOPERS CORNER (HTTPS://ANALYTICSINDIAMAG.COM/CATEGORY/DEVELOPERS CORNER/)

Demonstration Of What-If Tool For Machine Learning Model Investigation



BY BHOOMIKA MADHUKAR (HTTPS://ANALYTICSINDIAMAG.COM/AUTHOR/BHOOMIKA-MADHUKARANALYTICSINDIAMAG-COM/) 08/09/2020



(https://www.qpiai-explorer.tech/certification/?

<u>utm_source=aimagazine&utm_medium=banner&utm_campaign=preregistration)</u>

achine learning era has reached the stage of interpretability (intelligence-and-machine-learning/) where developing models and making predictions is simply not enough any more. To make a powerful impact and get good results on the data it is important to investigate and probe the dataset and the models. A good model investigation involves digging deep into the understanding of the model to find insights and inconsistencies in the developed model. This task usually involves writing a lot of custom functions. But, with tools like What-If, it makes the probing task very easy and saves time and efforts for programmers.

In this article we will learn about:

- 1. What is the What-If tool?
- 2. What are the features of this tool?
- 3. Walkthrough with a sample dataset.

What is the What-If tool?

(https://analyticsindiamag.com/)

What-If tool is a visualization tool that is designed to interactively probe the machine learning models. WIT allows users to understand machine learning models like <u>classification</u> (https://analyticsindiamag.com/product-sentiment-classification-weekend-hackathon-19/), regression (https://analyticsindiamag.com/a-beginners-guide-to-regression-techniques/) and deep neural networks (https://analyticsindiamag.com/deep-neural-networks-can-now-simulate-the-universe-but-scientists-cant-explain-how/) by providing methods to evaluate, analyse and compare the model. It is user friendly and can be used not only by developers but also by researchers and non-programmers very easily.

WIT was developed by Google under the <u>People+AI research (PAIR)</u> (https://analyticsindiamag.com/5-ways-in-which-ai-is-improving-accessibility-for-the-hearing-impaired/) program. It is open-source and brings together researchers across Google to study and redesign the ways people interact with AI systems.



(https://www.analytixlabs.co.in/)

What are the Features?

This tool provides multiple features and advantages for users to investigate the model.

Some of the features of using this are:

- 1. Visualizing the results of inference
- 2. Arranging the data according to the similarity
- 3. Editing data points to see how the model reacts to change
- 4. Comparing multiple machine learning models
- 5. Compare counterfactuals of the data
- 6. Implement confusion matrix and ROC curves
- 7. Use feature values to evaluate the model performance
- 8. Test algorithm constraints

WIT can be used with a Google Colab notebook or Jupyter notebook. It can also be used with Tensorflow Board.

Walkthrough with a Sample Dataset.

Let us take a sample dataset to understand the different features of WIT. I will choose the forest fire dataset available for download on Kaggle. You can click <u>here</u>

(https://www.kaggle.com/psvishnu/forestfire-impact-prediction-stats-and-ml) for downloading the dataset. The goal here is to predict the areas affected by forest fires given the temperature, month, amount of rain etc.

I will implement this tool on google collaboratory. Before we load the dataset and perform the processing, we will first install the WIT. To install this tool use,

!pip install witwidget

Importing the libraries and loading the data

```
import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import LabelEncoder
from sklearn.ensemble import GradientBoostingRegressor
from google.colab import drive
drive.mount('/content/gdrive')
df = pd.read_csv('/content/gdrive/My Drive/forestfires (2).csv')
df
```

what-if

Data splitting and pre-processing

```
features=df.drop('area',axis=1)
target=df[["area"]]
X_train, X_test, y_train, y_test = train_test_split(features, target,
test_size=0.2)
```

Once we have split the data, we can convert the columns month and day to categorical values using label encoder.

```
label_encoder=preprocessing.LabelEncoder()

X_train['month']=label_encoder.fit_transform((X_train['month']))

X_train['day']=label_encoder.fit_transform((X_train['day']))

X_test['month']=label_encoder.fit_transform((X_test['month']))

X_test['day']=label_encoder.fit_transform((X_test['day']))

X_train
```

machine learning

Now we can build our model. I will use sklearn ensemble model and implement the gradient boosting regression model.

Now that we have the model trained, we will write a function to predict the data since we need to use this for the widget.

```
def adjust_prediction(z):
   testing_data = pd.DataFrame(X_test, columns=X_test.columns.tolist())
   return reg.predict(testing_data)
```

Next, we will write the code to call the widget.

```
from witwidget.notebook.visualization impo [https://analytiesindiamag.com/)
from witwidget.notebook.visualization import WitWidget
num_data= 2000
tool height = 1000
test examples = np.hstack((X test[:num data].values,y test[:num data]))
config_builder = (WitConfigBuilder(test_examples.tolist(), X_test.columns.tolist()
  .set custom predict fn(adjust prediction)
  .set target feature('area')
  .set_model_type('regression'))
WitWidget(config builder, height=tool height)
```

This opens an interactive widget with two panels.

iupyter notebook

To the left, there is a panel for selecting multiple techniques to perform on the data and to the right is the data points.

As you can see on the right panel we have options to select features in the dataset along X-axis and Yaxis. I will set these values and check the graphs.

what-if tool

Here I have set FFMC along the X-axis and area as the target. Keep in mind that these points are displayed after the regression is performed.

Let us now explore each of the options provided to us.

Data exploration tab

Editing and viewing the dataset

You can select a random data point and highlight the point selected. You can also change the value of the datapoint and observe how the predictions change dynamically and immediately.

As you can see, changing the values changes the predicted outcomes. You can change multiple values and experiment with the model behaviour.

Finding Counterfactuals

Another way to understand the behaviour of a model is to use counterfactuals. Counterfactuals are slight changes made that can cause a model to flip its decision.

By clicking on the slide button shown below we can identify the counterfactual which gets highlighted in green.



(https://an alyticsindi amag.com/

8-best-

free-

resources-

to-learndeepDEVELOPERS CORNER (HTTPS://ANALYTICSINDIAMAG.COM/CATEGORY/DEVELOPERS_CORNER/)

8 Best Free Resources To Learn Deep Reinforcement Learning Using TensorFlow (https://analyticsindiamag.com/8-best-free-resources-tolearn-deep-reinforcement-learning-using-tensorflow/)

reinforcem
entlearningusingtensorflow
//

jupyter notebook

Partial Dependence Plots

This plot shows the effects that the features have on the trained machine learning model.

As shown below, we can see the inference of all the features with the target value.

jupyter notebook

Performance Tab

This tab allows us to look at the overall model performance. You can evaluate the model performance with respect to one feature or more than the one feature. There are multiple options available for analysis of the performance.

I have selected two features FFMC and temp against the area to understand performance using mean error.

If multiple training models are used their performance can be evaluated here.

Features Tab

The features tab is used to get the statistics of each feature in the dataset. It displays the data in the form of histograms or quantile charts.

The tab also enables us to look into the distribution of values for each feature in the dataset.

It also highlights the features that are most non-uniform in comparison to the other features in the dataset.

what-if tool

Identifying non-uniformity is a good way to reduce bias in the model.

Conclusion

WIT is a very useful tool for analysis of model performance. Ability to inspect models in a simple nocode environment will be of great help especially in the business perspective.

It also gives insights to factors beyond training the model like understanding why and how that model was created and how the dataset is fitting in the model.

What Do You Think?



Add a comment...



Bhavesh Wadibhasme

def adjust_prediction(z):

testing_data = pd.DataFrame(X_test, columns=X_test.columns.tolist()) return reg.predict(testing_data)

Here in this function insted of reg.predict() there should be reg_mod.predict().

Like · Reply · 24w

Facebook Comments Dive

Subscribe to our Newsletter

Get the latest updates and relevant offers by sharing your email.

ENTER YOUR EMAIL

SUBSCRIBE NOW

Join Our Telegram Group. Be part of an engaging online community. <u>Join Here (https://t.me/joinchat/NJLxnhZB7GkX3CPvjs9QGQ)</u>.



BHOOMIKA MADHUKAR (HTTPS://ANALYTICSINDIAMAG.COM/AUTHOR/BHOOMIKA-MADHUKARANALYTICSINDIAMAG-COM/)

I am an aspiring data scientist with a passion for teaching. I am a computer science graduate from Dayananda Sagar Institute. I have experience in building models in deep learning and reinforcement learning. My goal is to use AI in the field of education to make learning meaningful for everyone.

- f share: (https://www.facebook.com/sharer.php?u=https://analyticsindiamag.com/demonstration-of-what-if-tool-for-machine-learnin
 investigation/)
 - (https://twitter.com/intent/tweet?text=http://Demonstration%20Of%20What-
- **▼** TWEETIF%20Tool%20For%20Machine%20Learning%20Model%20Investigation&via=Analyticsindiam&url=https://analyticsindiamag.uof-what-if-tool-for-machine-learning-model-investigation/)
- in(https://www.linkedin.com/cws/share?url=https://analyticsindiamag.com/demonstration-of-what-if-tool-for-machine-learning-mc(https://wa.me/?text=http://Demonstration%200f%20What-
- ©If%20Tool%20For%20Machine%20Learning%20Model%20Investigation%20https://analyticsindiamag.com/demonstration-of-what (mailtne/subject-http://demonstration%20of%20What-
- If%20Tool%20For%20Machine%20Learning%20Model%20Investigation&body=http://Demonstration%20Of%20What-
- If%20Tool%20For%20Machine%20Learning%20Model%20Investigation%20https://analyticsindiamag.com/demonstration-of-what https://t.mm/share/will&textrabites://bernonstration%20Of%20What-
- ☐If%20Tool%20For%20Machine%20Learning%20Model%20Investigation&url=https://analyticsindiamag.com/demonstration-of-whatmachine-learning-model-investigation//popout?v=2&title=http://Demonstration%20Of%20What-
- ■If%20Tool%20For%20Machine%20Learning%20Model%20Investigation&url=https://analyticsindiamag.com/demonstration-of-what machine-learning-model-investigation/)



(https://ad.doubleclick.net/ddm/clk/491374101;298169448;q)



(https://business.louisville.edu//learnmore/UofLMSBA/?

utm_campaign=MSBA&utm_source=analyticsindia&utm_medium=display&utm_keyword=analyticsindia&utm_content=GetPaid)

OUR UPCOMING EVENTS

SKILLUP 2021 | Data Science Education Fair | 22-23rd April | Register here>> (https://skillup.analyticsindiasummit.com/) <u>Rising 2021 (https://rising.analyticsindiasummit.com/)</u> | Women in AI Conference | May 21 & 22 | Virtual

RELATED POSTS

OPINIONS (HTTPS://ANALYTICSINDIAMAG.COM/CATEGORY/ARTICLES/)

Best Reliable Deep-Tech To Track Criminals (https://analyticsindiamag.com/best- (https://analyticsi reliable-deep-tech-to-trackcriminals/)



reliable-deeptech-to-trackcriminals/)

19/03/2021 · 3 MINS READ

Does ML Hold The Promise To

Environmental Justice?

(https://analyticsindiamag.com/does-(https://analyticsi ml-hold-the-promise-toenvironmental-justice/)

ml-hold-thepromise-toenvironmentaljustice/)

14/03/2021 · 3 MINS READ

<u>DEVELOPERS CORNER</u>
(HTTPS://ANALYTICSINDI<u>AMAG.COM/CATEGORY/DEVELOPERS CORNER/)</u>

What is Trax and How is it a **Better Framework for Advanced Deep Learning?**

(https://analyticsindiamag.com/what-how-is-it-ais-trax-and-how-is-it-a-betterframework-for-advanced-deeplearning/)

(https://analytics is-trax-andbetterframework-foradvanced-deep-<u>learning/)</u>

14/03/2021 · 5 MINS READ

<u>DEVELOPERS CORNER</u> (HTTPS://ANALYTICS<u>INDIAMAG.COM/CATEGORY/DEVELOPERS_CORNER/)</u>

Guide to TensorFlow Extended(TFX): End-to-End **Platform for Deploying Production ML Pipelines**

(https://analytics to-tensorflowextendedtfx-

(https://analyticsindiamag.com/guide matform-forto-tensorflow-extendedtfx-endto-end-platform-for-deploying-<u>production-ml-pipelines/)</u>

platform-fordeployingproduction-mlpipelines/)

14/03/2021 · 5 MINS READ

OPINIONS (HTTPS://ANALYTICSINDIAMAG.COM/CATEGORY/ARTICLES/)

The Struggle For Data Privacy (https://analyticsindiamag.com/the-<u>struggle-for-data-privacy/)</u>

11/03/2021 · 3 MINS READ

(https://analytics struggle-fordata-privacy/)

<u>DEVELOPERS CORNER</u>
744745-77ANALYTICSINDIAM<u>AG.COM/CATEGORY/DEVELOPERS CORNER/)</u>

A Rrief Overview of OnenChat·



(https://analytics brief-overview-

of-openchatopen-source-

chattingframework-for-

generativemodels/)

BRAND PAGES

Career Center

LISTS

s=10+Emerging+Analytic

Framework for Generative

Models

<u>(https://analyticsindiamag.com/a-</u>

brief-overview-of-openchat-

open-source-chatting-

framework-for-generative-

OUR CONFERENCES

models/)

10/03/2021 · 3 MINS READ

OUR BRANDS

CONNECT WITH

MachineHack - ML Documentary - The Intel AI Hub **Academic Rankings** About Us <u>Cypher</u> (https://analyticsindiamag.(https://analyticsindiama <u>Hackathons</u> Transition Cost (https://analyticsindiamag/https://www.machinehack/https://www.analyticsind <u>| (https://www.youtube.com.on-intel/)</u> rankings/) v=7pvGJbzTTWk&list=PL9Kc1zSa46OzMfxoI1SJZOGpN_p37ivbd)
Best Firms To Work Advertise The MachineCon AIM Research (https://analyticsindiamag/https://aimresearch.ai/) (https://themachinecon.com/b Series - The ASSOCIATION OF with-us/) Dating Scientists (https://analyticsindiama Machine Learning **AIM Recruits** (https://www.youtube.com&cateNPTISTS best-firms-in-india-(https://recruits.analyticsir.dexelopers.Symmit Weekly Newsletter (http://mlds.analyticsindiav=WQbKblRKQsk&list=PL9Kc1zSa460xzJqQEJafor-data-scientists-(https://analyticsindiamagazine.substack.com/) qI55CtLZxFl2v) to-work-for-2021/) **Chartered Data** T<u>he Rising</u> Write for us Scientist(TM) (https://analyticsindiamag.AWARDS-(https://rising.analyticsind@dcaststrcSinylated Top Leaders (https://www.adasci.org/cd/https://analyticsindiama Reality for-us/) Careers (https://plugin.analyticsin/v=gvZfaeVVbGE&list=PL9KcisSa46QwqKw91j8W6vZ-(https://www.linkedin.com/conipany/analytics-(https://www.youtube.com/watch? Machine s=most+influential&cate Data Scientists (https://www.adasci.org/la(https://analyticsindiama 40 under 40 Data india-magazine/jobs/) **EVENTS** Analytics India Guru s=top+10+data+scientists **Scientists Continuous Learning** Contact Us (https://www.youtube.com/https://www.adasci.org/continuous_Startups (<u>https://mlds.analyticsindiasummit.com/awards/)</u> (https://analyticsindiamag <u>v=oQDZMdeyzgw&list=PL9learning/)</u>)yc1zxdtKwCO2S^j **AIM Custom Events** (https://analyticsindiama Data Science us/) (https://analyticsindiamag.com/powentious Geek

OUR VIDEOS

(https://www.analyticsindievents/aim-custom-award(https://www.youtube.com(https://www.adasci.org/career_award(https://www.youtube.com(https://www.adasci.org/career_award(https://www.youtube.com(https://www.adasci.org/career_award(https://www.youtube.com(https://www.adasci.org/career_award(https://www.youtube.com(https://www.adasci.org/career_award(https://www.youtube.com(https://www.adasci.org/career_award(https://www.youtube.com(https://www.adasci.org/career_award(https://www.youtube.com(https://www.adasci.org/career_award(https://www.youtube.com(https://www.adasci.org/career_award(https://www.youtube.com(https://www.adasci.org/career_award(https://www.youtube.com(https://www.adasci.org/career_award(https://www.adasci.org/career_award(https://www.youtube.com(https://www.adasci.org/career_award(https://www.youtube.com(https://www.adasci.org/career_award(https://www.youtube.com(https://www.adasci.org/career_award(https://www.youtube.com(https://www.adasci.org/career_award(https://www.youtube.com(https://www.adasci.org/career_award(https://www.youtube.com(https://www.adasci.org/career_award(https://www.youtube.com(https://www.adasci.org/career_award(https://www.youtube.com(https://www.adasci.org)) events/) v=Q7d1UR_PRGg&list=PL9center/)60xcRA2Y30FCFwj **MENTORSHIP**

(https://analyticsindiama Women in AI AIM Virtual Deeper Insights with s=top+trends&category=. <u>Leadership</u> <u>Membership</u>

(https://rising.analyticsind(https://analyticsindiamag_com/our-AIM Mentorship Circle (https://www.adasci.org/members-

(https://analyticsindiamag.com/mentorshipevents/aim-virtual/) (https://www.youtube.combenefits/)

circle/) v=Z3Z7Riw1G9o&list=PL9Kc1zSa46Oyv8tAFzC22cuLXNRUZvXAG)

Assisted Mentoring Curiosum – AI (https://analyticsindiamag.com/mentorship-Storytelling

circle/assisted-(https://www.youtube.com/watch?

mentoring/) v=AfsqH5EzjIg&list=PL9Kc1zSa46OzOCjo2YUNym6thlDDoFozc)

ABOUT US(HTTPS://ANALYTICSINDIAMAG.COM/ABOUT/)

Excellence

ADVERTISE(HTTPS://ANALYTICSINDIAMAG.COM/ADVERTISE-WITH-US/)

WRITE FOR US(HTTPS://ANALYTICSINDIAMAG.COM/WRITE-FOR-US/)

<u>COPYRIGHT(HTTPS://ANALYTICSINDIAMAG.COM/COPYRI<mark>GHT-TRADHMA</mark>(lastps://analyticsindiamag.com/)</u>



PRIVACY(HTTPS://ANALYTICSINDIAMAG.COM/PRIVAC

TERMS OF USE(HTTPS://ANALYTICSINDIAMAG.COM/TERMS-USE/)

 $\underline{CONTACT\ US(HTTPS://ANALYTICSINDIAMAG.COM/CONTACT-US/)}$

COPYRIGHT ANALYTICS INDIA MAGAZINE PVT LTD

(HTTPS://MEDIUM.COM/ANALYTICS-INDIA-MAGAZINE) in
(HTTPS://WWW.LINKEDIN.COM/COMPANY/ANALYTICS-INDIA-MAGAZINE) (HTTPS://T.ME/NJLXNHZB7GKX3CPVJS9QGQ)