|  |  |
| --- | --- |
| **Server & Cloud Security**  Diploma in CSF/IT  Oct 2022 | Week 13 |
| Practical |
| **Practical: Splunk Forwarder & Machine Learning Demos** | |

## Learning Objectives

You will learn how to configure and use the Splunk Forwarder to perform security monitoring, and analyze the data using machine learning algorithms.

**Part 1: Installation & Preparation**

**You have installed the Spunk Enterprise on your MS1 server in the previous practical.**

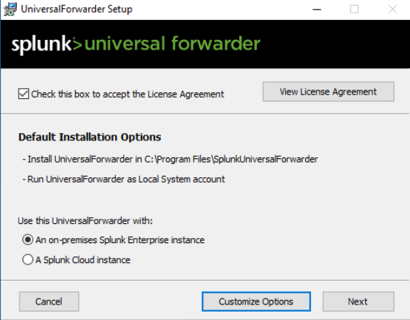
For this practical you will install Splunk Forwarder on **DC1**.

Firstly, download the installer from PoliteMall, or from the following Google Drive link.

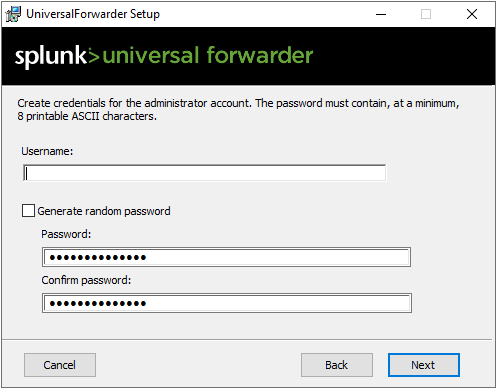
Google Drive:

<https://drive.google.com/file/d/1jAe-FEE0Q1mbeGDnPtm2BfnZtrF2IBe1/view?usp=sharing>

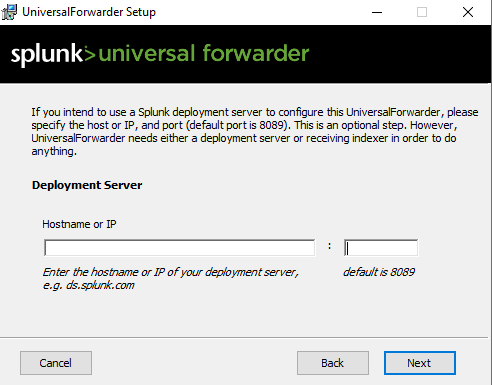
1. Double-click the installer and agree the license agreement.



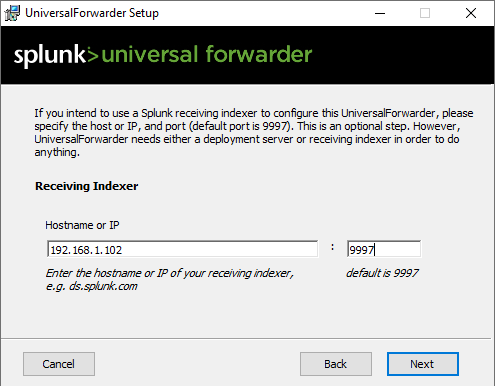
1. Click **Next**.
2. Use the same admin username and password as you created for Splunk Enterprise in the last practical.



1. Leave the deployment server blank.

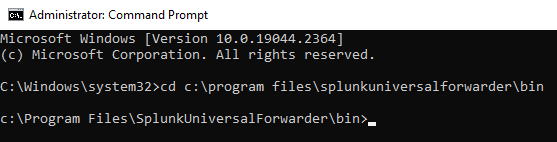


1. Set the receiving indexer using the IP address of your MS1 server and use the default port 9997.

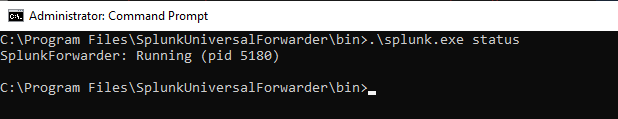


1. Click Next till the installation finishes.
2. To ensure if Splunk is working, open cmd as administrator and type the following command:

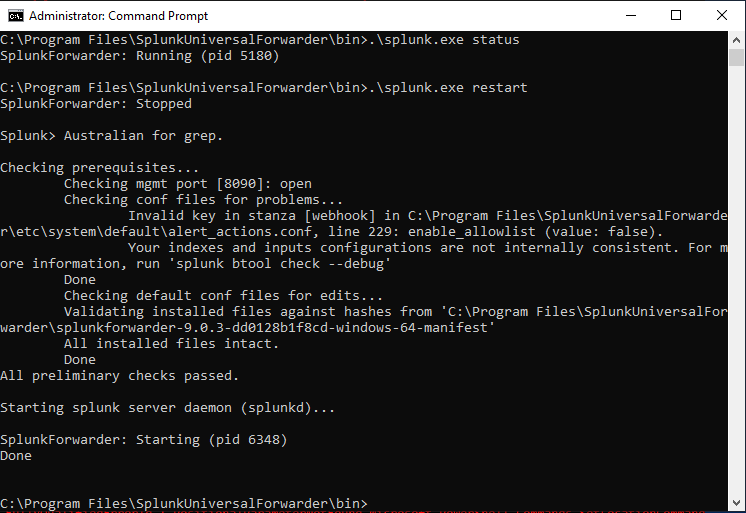
***cd C*:\*Program Files\SplunkUniversalForwarder\bin***



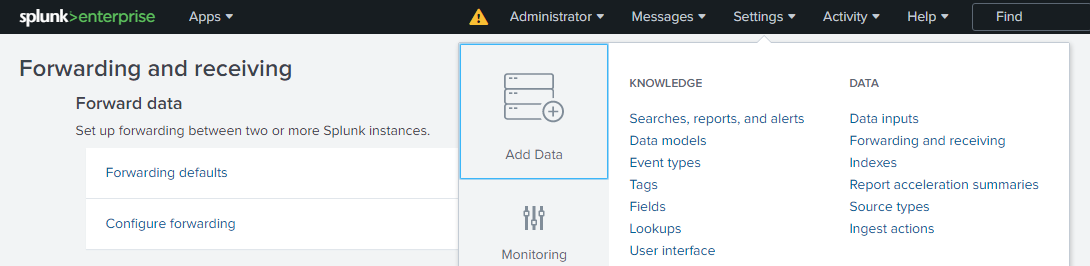
1. Type **.\splunk.exe** **status** to verify if the Splunk Forwarder is running.



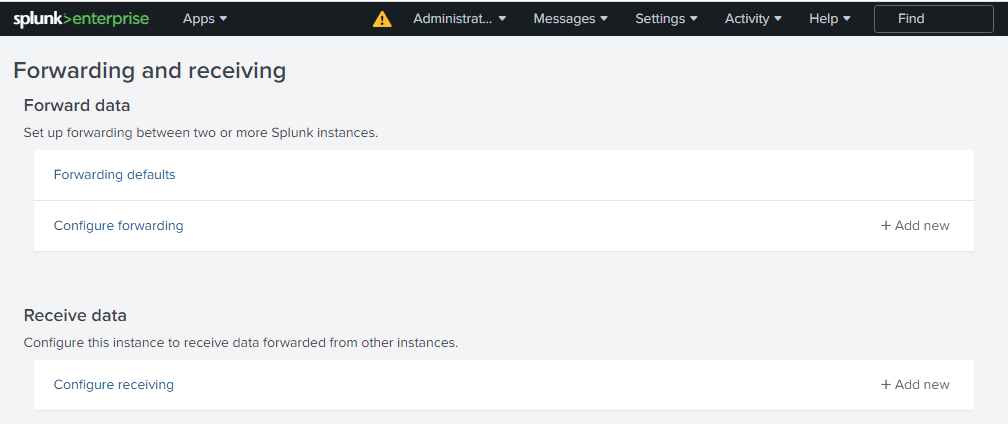
1. **[Optional]** If the forwarder is not running, type **.\splunk.exe restart** to restart it.



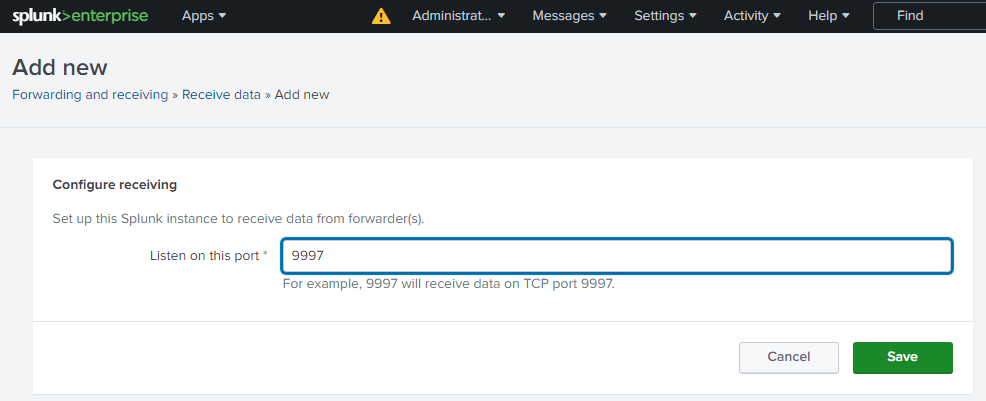
1. Switch to MS1, open your browser and use the URL to access Splunk Web <http://localhost:8000>.
2. Click “Settings” and then select “Forwarding and receiving”.



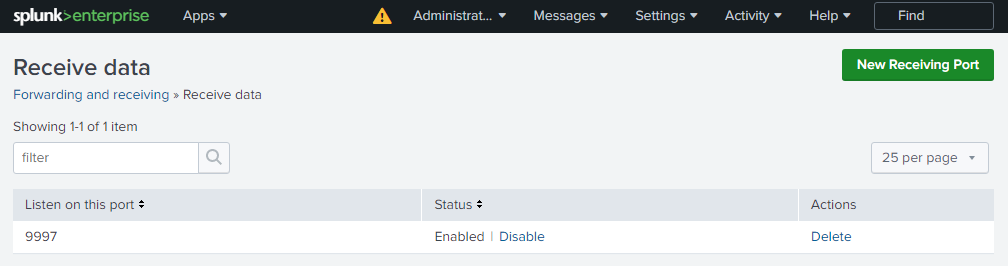
1. Select “Add new” to configure receiving.



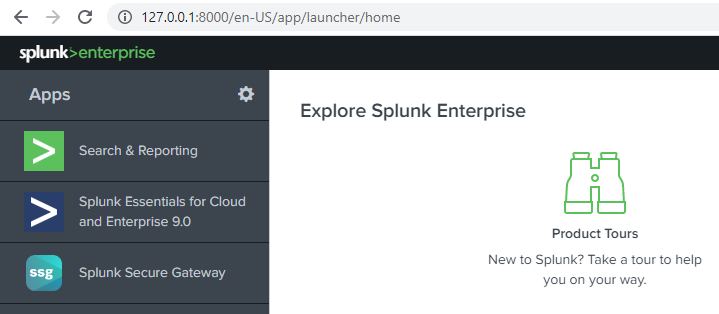
1. Add port 9997 and click save.



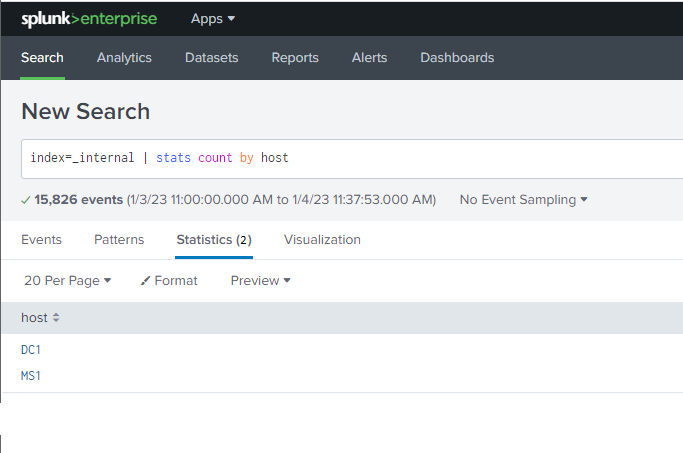
1. Now your Splunk Enterprise is listening on the port 9997.



1. Return to the main page of the Splunk Enterprise and select “Search & Reporting”.



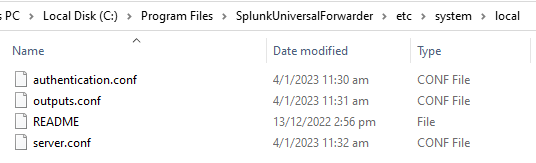
1. Under the search tab search for **index=\_internal | stats count by host** to confirm that the forwarder is installed.

****

**Part 2: Forwarding Windows Events**

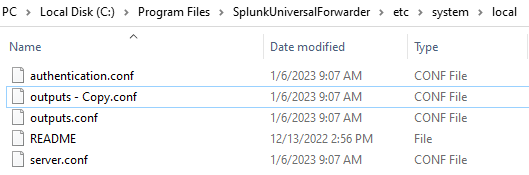
1. Switch to DC1 where you installed Splunk Forwarder and Access the following path:

C:\Program Files\SplunkUniversalForwarder\etc\system\local

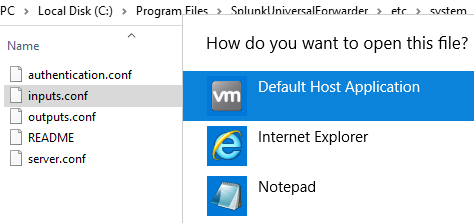


Note that after you install a fresh new Splunk Forwarder, there is no **inputs.conf** file in the folder.

1. Make a copy of the outputs.conf file.



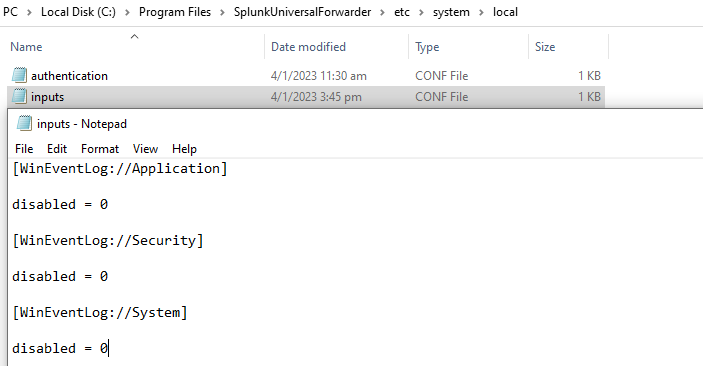
1. Rename it as inputs.conf and then use Notepad to open it.



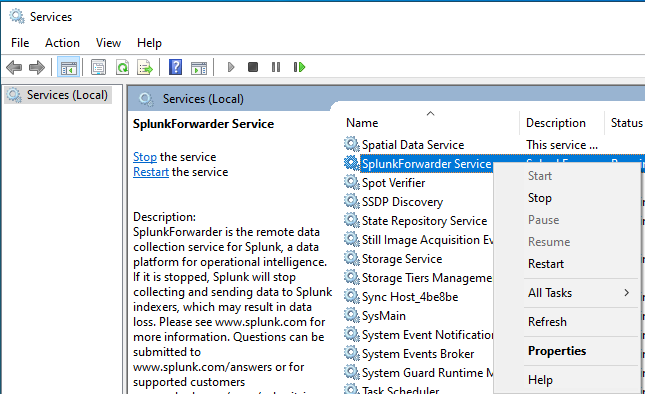
1. Delete the original contents and paste the below lines into the config file then save it.

|  |
| --- |
| [WinEventLog://Application]  disabled = 0  [WinEventLog://Security]  disabled = 0  [WinEventLog://System]  disabled = 0 |

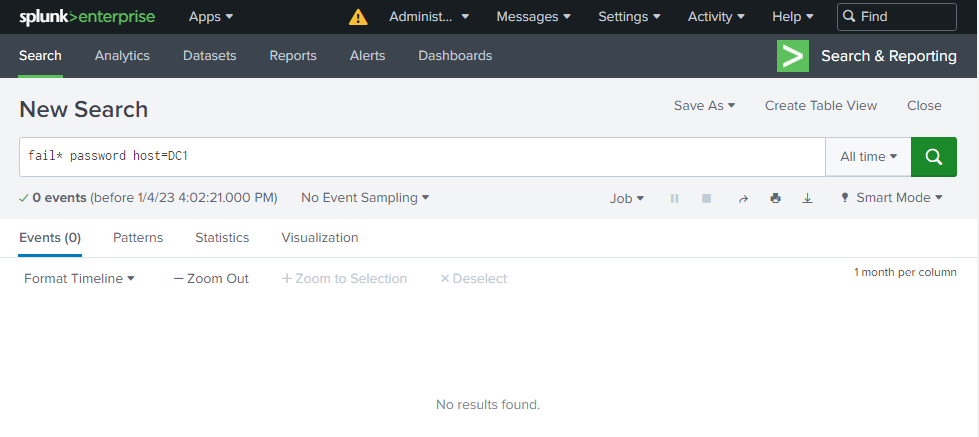
1. Make sure that you have the inputs.conf file and the contents are as shown.



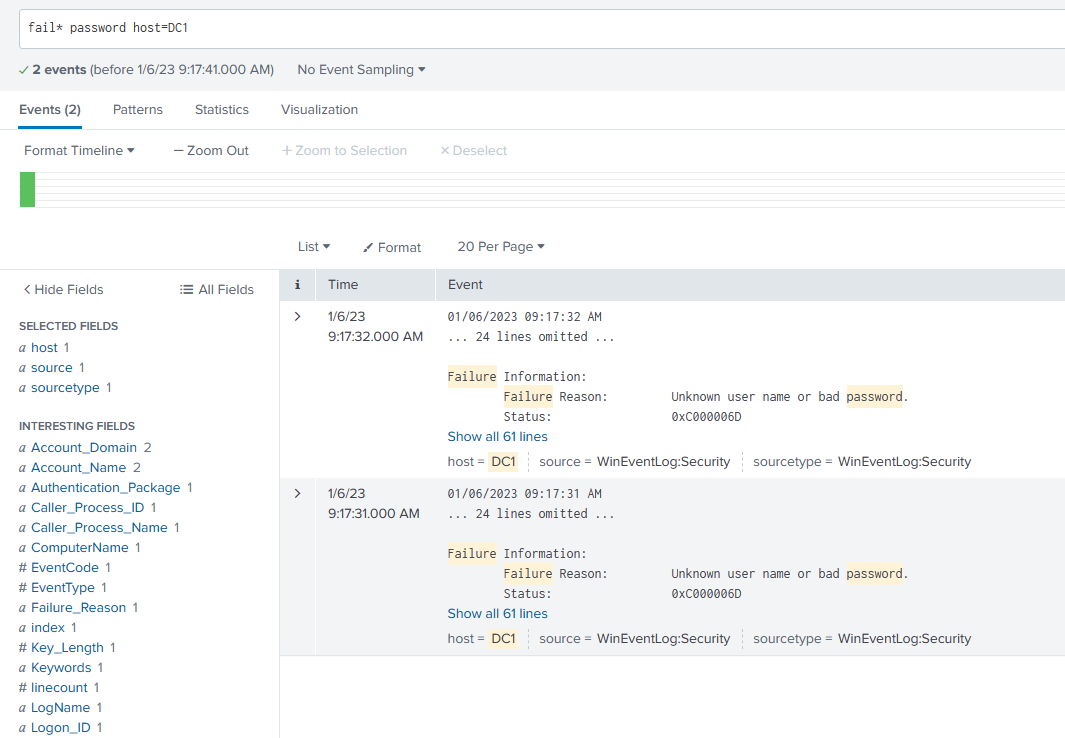
1. Press the **Start** [Image titled Windowsstart.png](https://www.wikihow.com/Open-Windows-Services#/Image:Windowsstart.png) button and type **services**.
2. Right click the SplunkForwarder Service and click Restart.



1. Now your Splunk Forwarder is able to forward the Windows Application log, Security log and System log to the Splunk Enterprise installed on MS1.
2. Start a Search and key in **fail\* password host=DC1**, make sure the time range is set to “All Times”. It’s correct to have no results found.

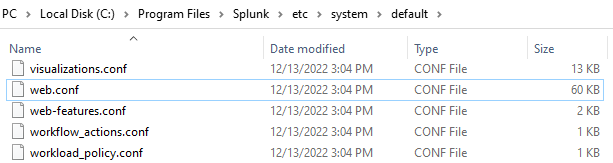


1. Sign out from DC1 and try to sign in with wrong password for **2** times.
2. On MS1 Splunk Enterprise, do the search as shown in step 1 again. You should be able to see the security event log forwarded to MS1 from DC1. There should be a total of 2 events for authentication failure.

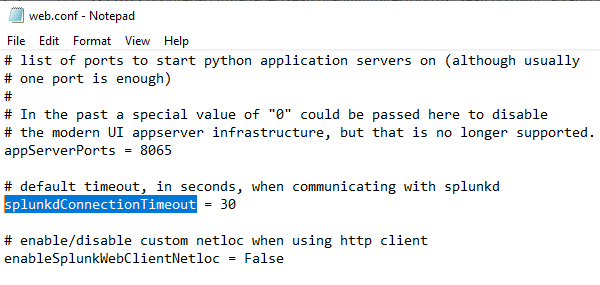


**Part 3: Machine Learning Toolkit Installation**

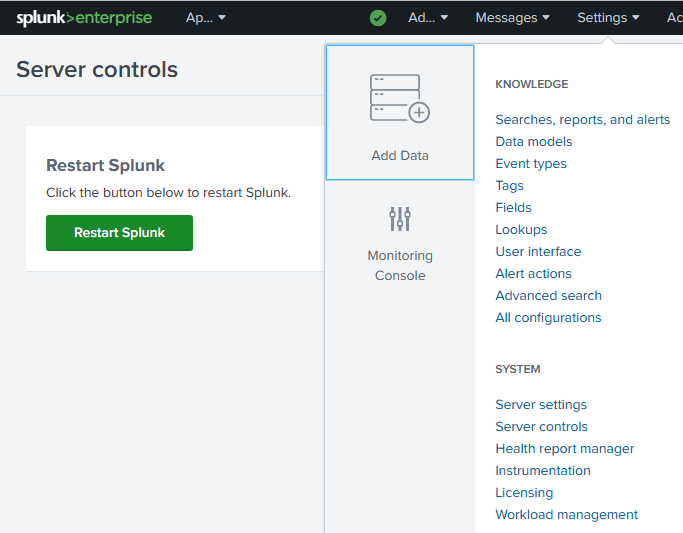
1. On MS1, access the following path of the Splunk Enterprise folder.



1. Open the web.conf file using Notepad. Search for the keywork **splunkdConnectionTimeout** to locate the line, then change the default timeout from 30 to 300. This is to make sure the Splunk server would have enough waiting time when you install the new App components.



1. Click Settings and then Server Controls, restart Splunk and then log in again.

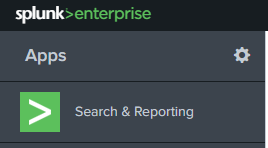


1. Download the “splunk-machine-learning-toolkit\_533.tgz” and “python-for-scientific-computing-for-windows-64-bit\_302.tgz” packages from Politemall or the below Google Drive links. Place them into your MS1.

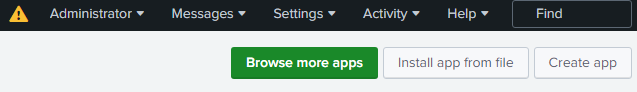
<https://drive.google.com/file/d/1gD7gS1XJBypWFHiioTFT2ih6FsGXlYUW/view?usp=sharing>

<https://drive.google.com/file/d/1hrALxOfjln0CW-ERYq15uTjz07ZsWvaJ/view?usp=sharing>

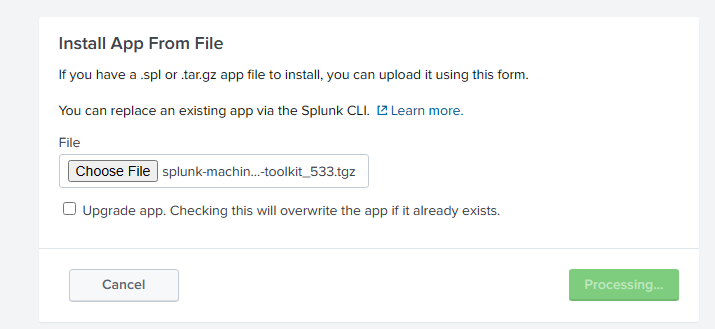
1. Click the below icon to access the managed Apps.



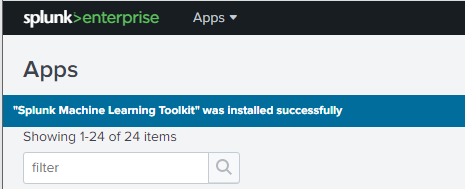
1. Select “Install app from file”.



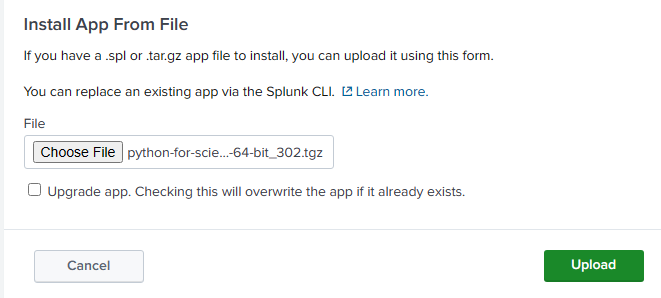
1. Choose the “splunk-machine-learning-toolkit\_533.tgz” that you have downloaded to install it as an App of Splunk.



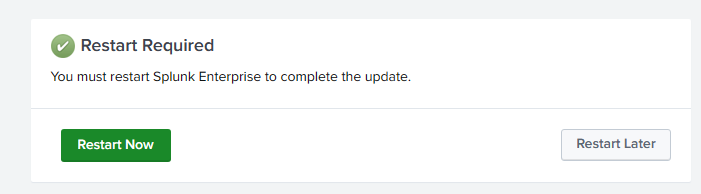
1. Confirm that the Machine Learning Toolkit is installed successfully.



1. Select “Install app from file” again and repeat the same steps to install “python-for-scientific-computing-for-windows-64-bit\_302.tgz”.

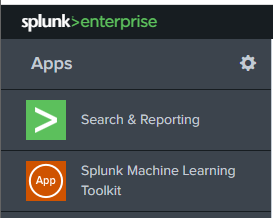


1. After installation, you need restart the Splunk again.

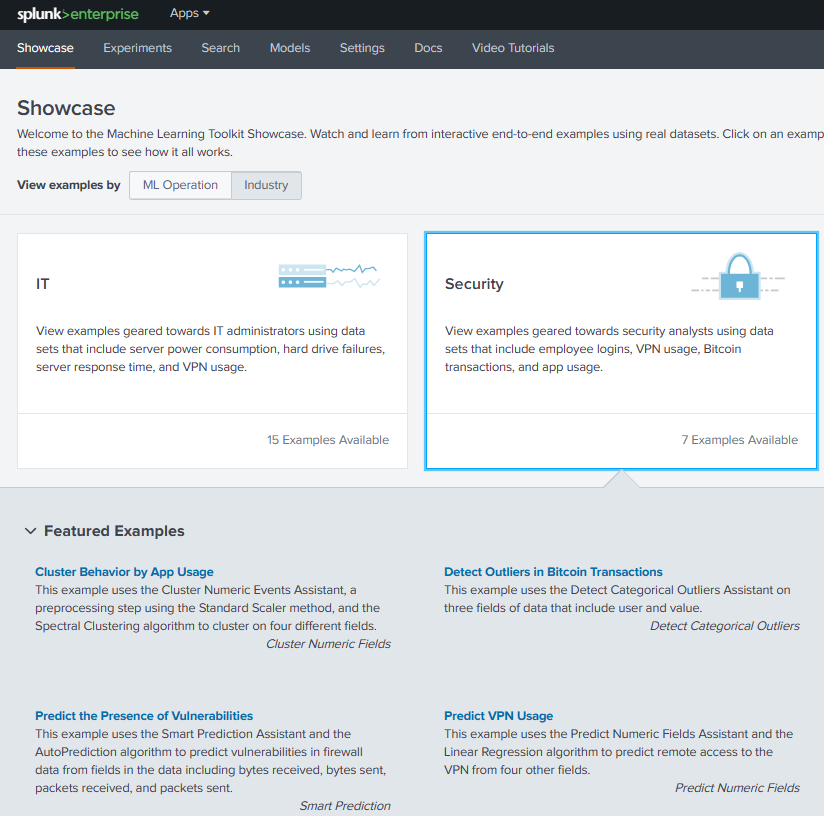


**Part 4: Machine Learning Toolkit Demo 1**

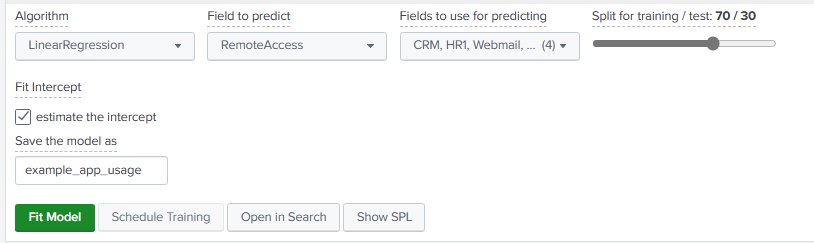
1. Login to Splunk Enterprise after restart. You should be able to see the new app – Splunk Machine Learning Toolkit installed, click and open it.



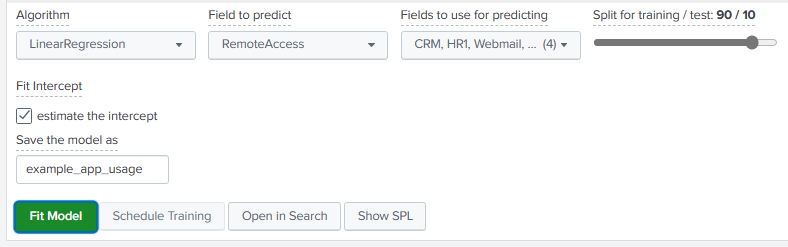
1. On the Showcase tab, choose Industry\Security and then select “Predict VPN Usage”.



1. The data set: app\_usage.csv is auto loaded and the result is automatically populated too.
2. Note that the algorithm used is Linear Regression. As a recap from Week 12 lecture, regression algorithm belongs to supervised learning which requires labelled data to train the algorithm. Thus 70% of the events are used for training and the rest 30% are using for testing.



1. Observe the generated charts to understand the accuracy of the prediction.
2. Note that the R2 value for this model is 0.9179. R2 is always between 0~1, a better module has an R2 value closer to 1. In a simple word, R2 shows how well the data fit the regression model.
3. Split the training / test data as 90/10 so that you have more labeled data to train the algorithm.

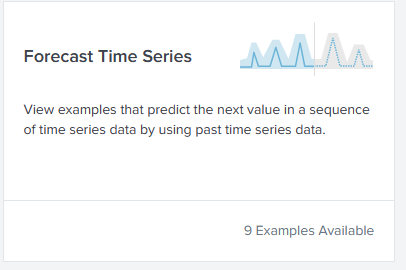
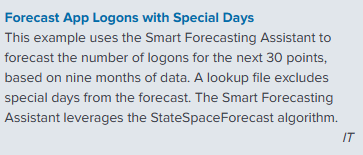


1. Discuss what is the R2 value now, and what can you interpret from this value?

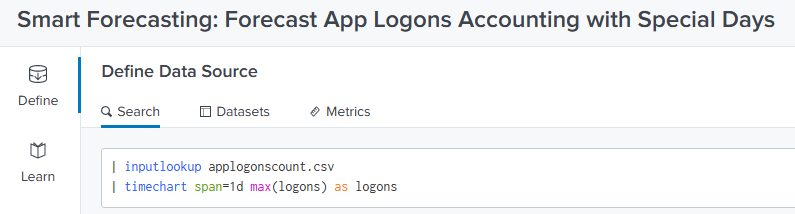
|  |
| --- |
| **0.9406. The model has a better data fit.** |

**Part 5: Machine Learning Toolkit Demo 2**

1. Return to the Showcase tab, choose Forecast Time Series and then select “Forecast App Logons with Special Days”.

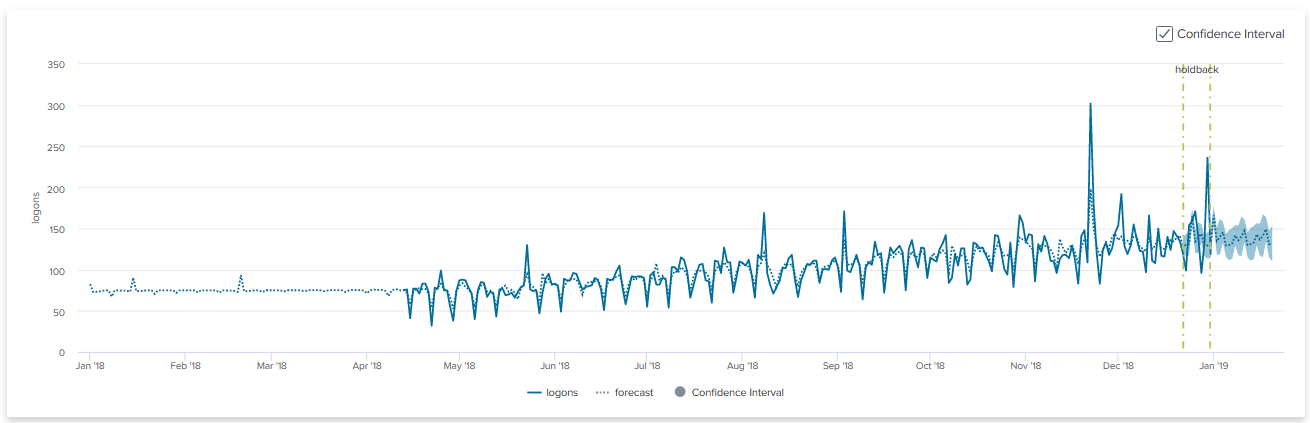
1. The Forecast Time Series is used to predict the next value in a sequence of time series data. The result includes both the predicted value and a measure of the uncertainty of that prediction.
2. **[Optional]** The algorithm used in this demo is Kalman Filter algorithm.
3. To define Data Source, the query is auto filled, click search to start.



1. Click “Learn”, make sure to set the following parameters and click “Forecast”. Mouse over to the “question marks icons to understand all the parameters.

|  |  |
| --- | --- |
|  | The filed “logons” contains null values because for some specific dates, there was no logon event happened.  The algorithm is able to impute/estimate a proper value automatically. |
| Fields are the columns in the dataset. In this example, we select “logons” to forecast the future number of logons. |
| Holdback: |
| Future Timespan: |
| Confidence Interval: |
| Special days: |
| Period: |

1. You will have an evaluation results similar to the below. Note that the holdback period contains the 10 data points (with known values) used to compare and evaluate the forecast values.



== End of Practical ==