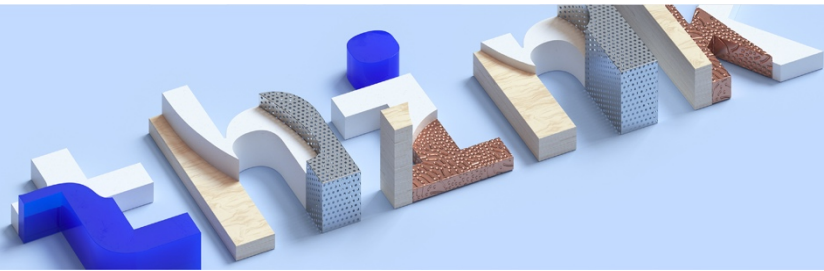


think 2018



Lab Center – Hands-on Lab

Session 4874A

Visualizing Deployed Models with R Shiny and Watson Machine Learning

Greg Filla, IBM, gfilla@us.ibm.com

Table of Contents

Disclaimer	3
Sign up for IBM Watson Studio using your IBM Cloud Account.....	5
Working with Watson Studio	6
Create a project	6
Add data assets to project	6
Building a flow in Modeler	7
Adding data using the flow	7
Preparing data for modeling	7
Training the model	8
Saving the model to the project	10
Creating a deployment for the model	11
Building your Shiny Dashboard	12
We Value Your Feedback!	13

Disclaimer

IBM's statements regarding its plans, directions, and intent are subject to change or withdrawal without notice at IBM's sole discretion. Information regarding potential future products is intended to outline our general product direction and it should not be relied on in making a purchasing decision.

The information mentioned regarding potential future products is not a commitment, promise, or legal obligation to deliver any material, code or functionality. Information about potential future products may not be incorporated into any contract.

The development, release, and timing of any future features or functionality described for our products remains at our sole discretion I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve results like those stated here.

Information in these presentations (including information relating to products that have not yet been announced by IBM) has been reviewed for accuracy as of the date of initial publication and could include unintentional technical or typographical errors. IBM shall have no responsibility to update this information. **This document is distributed “as is” without any warranty, either express or implied. In no event, shall IBM be liable for any damage arising from the use of this information, including but not limited to, loss of data, business interruption, loss of profit or loss of opportunity.** IBM products and services are warranted per the terms and conditions of the agreements under which they are provided.

IBM products are manufactured from new parts or new and used parts.

In some cases, a product may not be new and may have been previously installed. Regardless, our warranty terms apply.”

Any statements regarding IBM's future direction, intent or product plans are subject to change or withdrawal without notice.

Performance data contained herein was generally obtained in controlled, isolated environments. Customer examples are presented as illustrations of how those customers have used IBM products and the results they may have achieved. Actual performance, cost, savings or other results in other operating environments may vary.

References in this document to IBM products, programs, or services does not imply that IBM intends to make such products, programs or services available in all countries in which IBM operates or does business.

Workshops, sessions and associated materials may have been prepared by independent session speakers, and do not necessarily reflect the views of IBM. All materials and discussions are provided for informational purposes only, and are neither intended to, nor shall

constitute legal or other guidance or advice to any individual participant or their specific situation.

It is the customer's responsibility to insure its own compliance with legal requirements and to obtain advice of competent legal counsel as to the identification and interpretation of any relevant laws and regulatory requirements that may affect the customer's business and any actions the customer may need to take to comply with such laws. IBM does not provide legal advice or represent or warrant that its services or products will ensure that the customer follows any law.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products about this publication and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products. IBM does not warrant the quality of any third-party products, or the ability of any such third-party products to interoperate with IBM's products. **IBM expressly disclaims all warranties, expressed or implied, including but not limited to, the implied warranties of merchantability and fitness for a purpose.**

The provision of the information contained herein is not intended to, and does not, grant any right or license under any IBM patents, copyrights, trademarks or other intellectual property right.

IBM, the IBM logo, ibm.com and [names of other referenced IBM products and services used in the presentation] are trademarks of International Business Machines Corporation, registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the Web at "Copyright and trademark information" at: www.ibm.com/legal/copytrade.shtml.

© 2018 International Business Machines Corporation. No part of this document may be reproduced or transmitted in any form without written permission from IBM.

U.S. Government Users Restricted Rights — use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM.

Sign up for IBM Watson Studio using your IBM Cloud Account

If you have not yet tried IBM Watson Studio using your IBM Cloud Account then follow the steps below. Otherwise jump to the next section.

1. Go to datascience.ibm.com
2. Click Sign Up

If you already have an IBM Cloud Account then click on 'Use it to sign up for IBM Watson'. Then click on the 'IBM Watson Studio' logo in the top left navigation bar, which will take you to the 'Welcome' page.

If you do not already have an IBM Cloud Account then do so through the 'Register for IBM Cloud' option.

1. Register using your email address.
2. Create your free IBM Cloud account and verify your email.
3. Once you log in to your IBM Cloud account after verification you will be taken to Watson Studio.

Working with Watson Studio

The purpose of this lab session is to learn about the machine learning lifecycle and gain skills in deploying and visualizing predictive models. We will start the lab by adding data from the Watson Studio community into a new project then quickly building a model using SPSS Modeler flows. Once we train a model, we can save it in the Watson Machine Learning (WML) repository where it can easily be deployed. Once we have the deployment we will create a dashboard using RStudio and Shiny.

Create a project

All analytical assets within Watson Studio are organized into 'Projects'. Projects are containers which can include data assets, Notebooks, Flows and models among other items. Projects can be personal or shared with a team of collaborators. It is a best practice to create a new project when you start a new analytics task.

1. From the 'Welcome' page of Watson Studio click on 'New project'.
2. Enter a project name, and select your target IBM Cloud Object Storage Instance.
3. If you do not have a Cloud Object Storage service defined then click on 'Add' and select the 'Lite' plan, click 'Create' and then 'Confirm', and click 'Refresh' to detect this as the defined compute engine, then click 'Create'.

Add data assets to project

Data assets are added to a project to make them available for any of the tools included in Watson Studio. Data assets added to a project are also available for any of the collaborators sharing that project to use. Examples of data assets are files, such as .csv or .json or database tables added through a data connection.

1. Click on the 'Community' tab in the top bar. From here you can see content in the Watson Studio Community, which includes data sets.
2. Navigate to the data set 'UCI: Adult - Predict income', then click on the add (+) icon on the community card for this data set, select your project and click 'Add'. Once this completes it will display 'Added'. This may take up to a minute.
3. Click on the 'View Project' link.
4. Within the project's 'Assets' tab you will see 'Data assets' shown including the one added above. Click on the 'UCI: Adult - Predict income' name to preview the data.
5. Return to the project by clicking on the project name in the navigation breadcrumb.

Building a flow in Modeler

1. From the navigation bar select 'Add to project' then 'Modeler flow'.
2. Provide a name and use the default runtime 'IBM SPSS Modeler', and click 'Create'.

Adding data using the flow

1. On the right side of the screen there is the list of the data assets which we added previously. Drag and drop the 'UCI: Adult - Predict income' data asset onto the canvas in the middle of the screen.

This dataset contains demographic information for the for adults including whether their salary is above or below \$50k annually. The original data source can be found here:

<https://archive.ics.uci.edu/ml/datasets/Adult>.

2. Right click on the 'UCI: Adult - Predict income.csv' node, or click on the three dots within the node, and select Preview to examine this data set.

As you can see this dataset contains a number of columns; including information about adults in the study like their age, education, marital status, etc. You can click the **Profile** and **Visualizations** tabs to explore the data more. Since this lab is more about understanding the full machine learning lifecycle we won't focus much on exploratory data analysis, but rather move forward to model building. Click the name of your flow in the breadcrumb to return to the flow editor.

Preparing data for modeling

1. Add a Filter node from the node palette on the left hand side. The Filter node is a 'Field Operations' node (or you can use the search bar).
2. Connect this node to the data set node then double click to open the properties on the right side.
3. Clicking the add columns button allows you to choose which fields you want to include for your model. Use the filtering available to select the following fields:

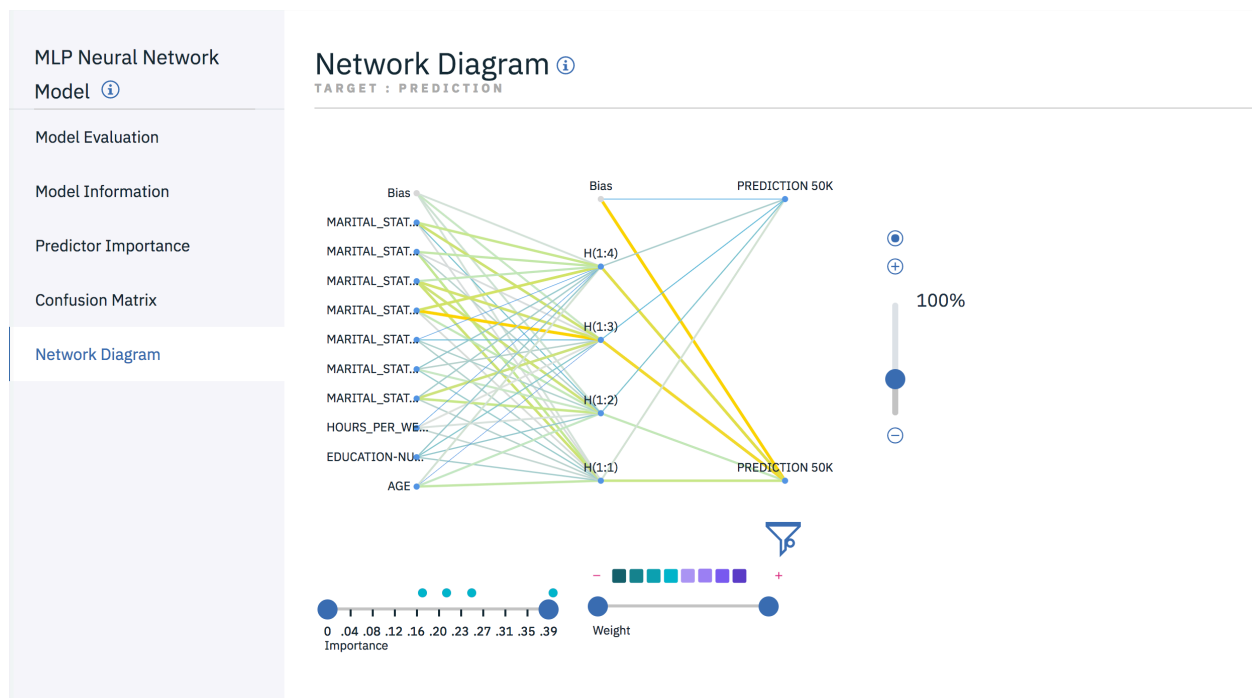
- Age
- Education-Num
- Marital_status
- Occupation
- Hours_per_week
- Prediction

4. Click **OK** then change the Mode to "Retain the selected fields..." then click **Save**

- Now add a Type node from the node palette on the left hand side. The Type node is a 'Field Operations' node (or you can use the search bar).
- Double click the Type node and click on 'Configure Types' then click 'Read Values'.
- Change the *Role* for Prediction to be **Target**, then click **OK** then **Save**
- This completes the minimal pre-processing we need to do before building a model. Since our target is discrete (either over 50k salary, or below) we need to choose an algorithm for classifying data.

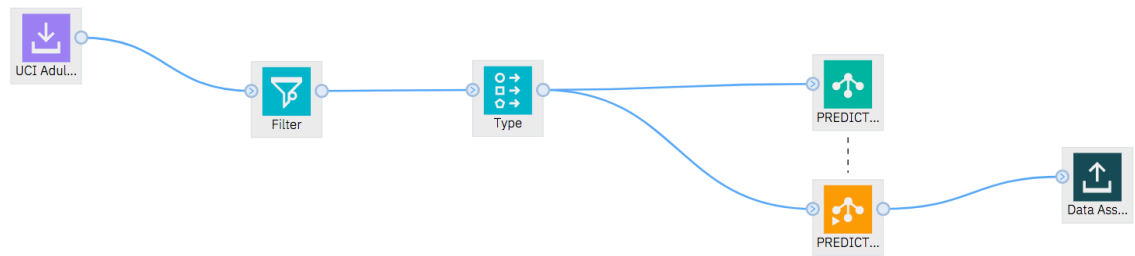
Training the model

- Go back to the palette and select a **Neural Network** model to add to the flow.
- Right click this node and click **Run** to begin model training
- To investigate this model, right the the model object (orange node) and click **View Model**



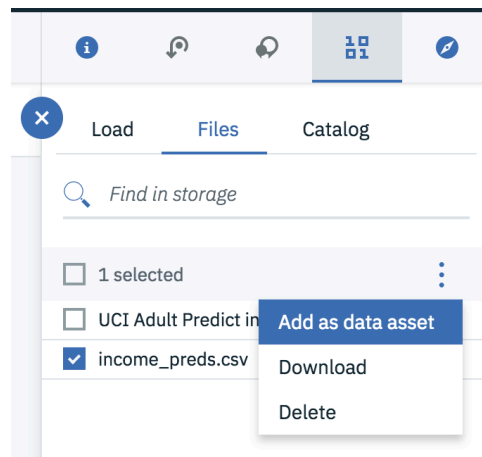
Now we have a trained model that can be deployed to Watson Machine Learning.

- From the Export section of the Palette, add a Data Asset Export node to the flow and connect it to the Neural Network model. Double click this and add a name for the file, then run the flow to train the model and export the scored data to Cloud Object Storage



Your flow should look something like this

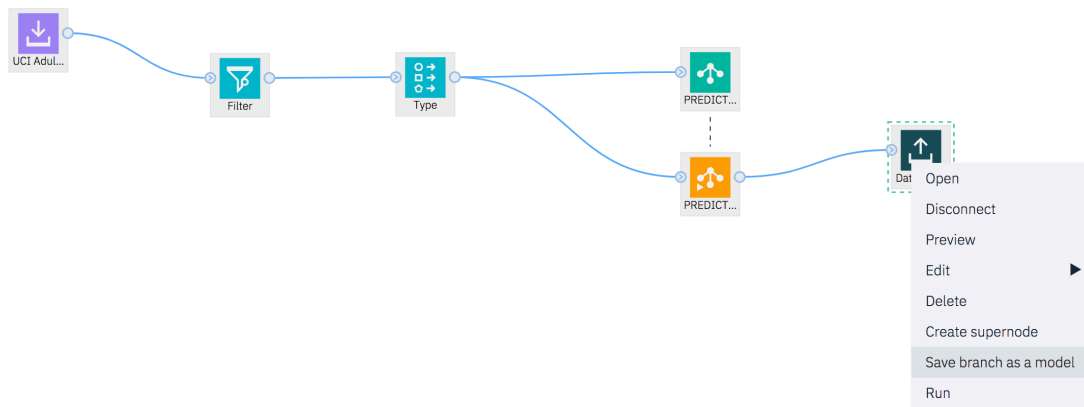
5. You can find this data asset in the right side panel of your project, where it can be added as a project asset if needed.



Saving the model to the project

To use the model for scoring we need to deploy it in the Watson Studio project. The first step in this process is defining the branch in the flow that will be scored with data.

1. To deploy this flow, right-click on the final node (“Data Asset Export”) and click “Save branch as a model”.



At this point you will be required to create a new Watson Machine Learning service instance to save your model to.

2. In the ‘Required Service Missing’ dialog click on the ‘Create a new Watson Machine Learning service instance’ link.
3. Select the ‘Lite’ plan then click ‘Create’.
4. In the ‘Confirm Creation’ dialog click ‘Confirm’.

You will be returned to the flow and from this point you can use the context menu on the table node click ‘Save as a model’ to use this newly defined Watson Machine Learning service instance.

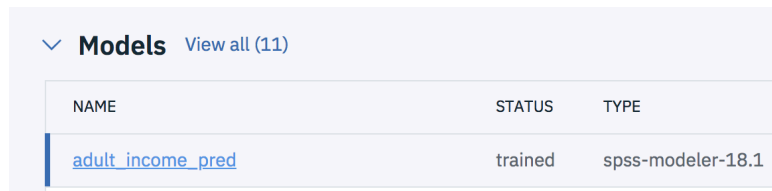
5. On the ‘Save Model’ page give the model a suitable name (such as ‘Income Model’) and make sure the Watson Machine Learning instance you created earlier is selected.
6. Then click ‘Save’.

The model has now been saved as an analytics asset in your Watson Studio project making it available for deployment for scoring.

Creating a deployment for the model

In order to use this model for our Shiny dashboard, we need to create a deployment so we can score new data with the model via REST API.

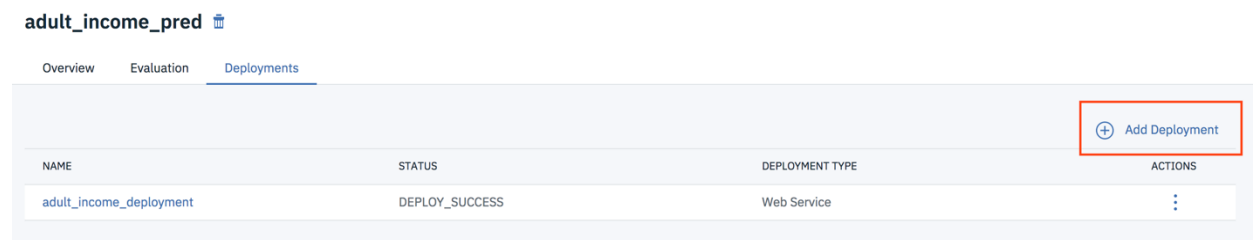
1. Go to the Models section of your project and click the name of the saved Model to see its details:



Models View all (11)

NAME	STATUS	TYPE
adult_income_pred	trained	spss-modeler-18.1

2. After drilling into the model page, go to the Deployments tab and add a deployment:



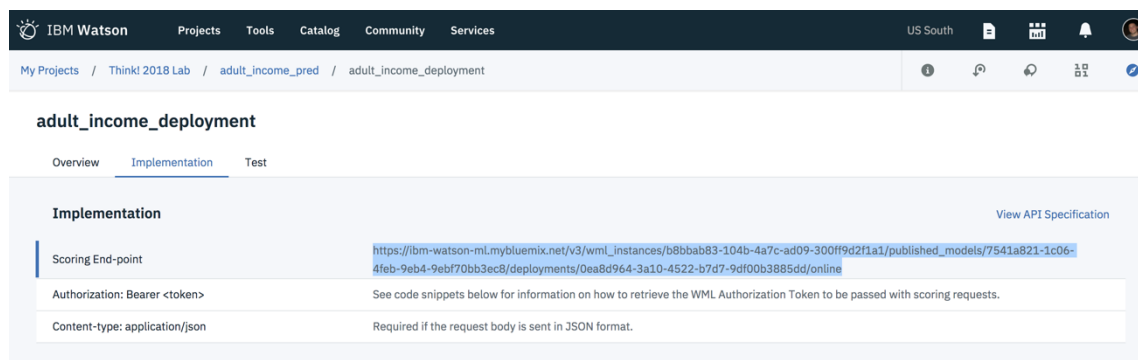
adult_income_pred

Overview Evaluation Deployments

+ Add Deployment

NAME	STATUS	DEPLOYMENT TYPE	ACTIONS
adult_income_deployment	DEPLOY_SUCCESS	Web Service	

3. After creating the deployment, you can view more information clicking on the name of the deployment, or clicking on the **Deployments** tab of the project to view all deployments
4. The important information we need from the deployment page is the scoring end-point – this is the url we will use for the dashboard code:



IBM Watson Projects Tools Catalog Community Services US South

My Projects / Think! 2018 Lab / adult_income_pred / adult_income_deployment

adult_income_deployment

Overview Implementation Test

Implementation View API Specification

Scoring End-point	https://ibm-watson-ml.mybluemix.net/v3/wml_instances/b8bbab83-104b-4a7c-ad09-300ff9d2f1a1/published_models/7541a821-1c06-4feb-9eb4-9ebf70bb3ec8/deployments/0ea8d964-3a10-4522-b7d7-9df00b3885dd/online
Authorization: Bearer <token>	See code snippets below for information on how to retrieve the WML Authorization Token to be passed with scoring requests.
Content-type: application/json	Required if the request body is sent in JSON format.

Building your Shiny Dashboard

To complete this lab, please navigate to <https://github.com/gfilla/PredictingIncomeDashboard>

This repository contains the following steps:

- Finding your Watson Machine Learning credentials
- Understanding how R Shiny and Flexdashboards work
- Understanding how to score deployed models via REST API
- Composing a dashboard that scores your deployed model

We Value Your Feedback!

- Don't forget to submit your Think 2018 session and speaker feedback! Your feedback is very important to us – we use it to continually improve the conference.
- Access the Think 2018 agenda tool to quickly submit your surveys from your smartphone, laptop or conference kiosk.