



Author

Rahul Sinha, MBA '20

Data Management and Business
Process Modeling

RECOMMENDATION ENGINE

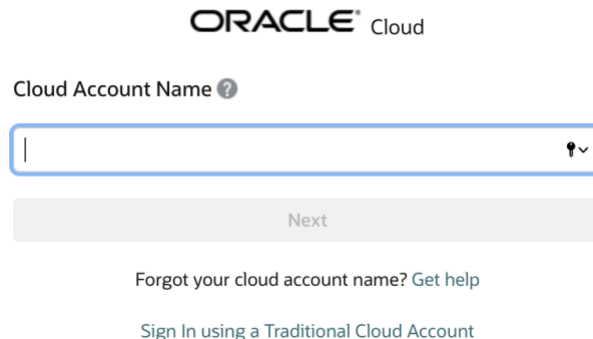
Recommendation Engine using Oracle Cloud

Problem Statement: After developing our Entity Relationship Diagrams, the required tables in an Oracle database, and a few report queries using Oracle's SQL Developer, we wanted to create for our client more value from our Data Management Solution. We realized that building a recommendation engine would help the client's sales team improve its cross-sell probability by helping the team make data-informed recommendations to their customers.

I researched some possible solutions, and realized that I could implement a recommendation engine using Oracle Cloud. Here is a quick run through the steps involved.

Part I: Setup Oracle Cloud and create OML user

1. Create an Oracle Cloud Free Tier Account
(<https://www.oracle.com/cloud/free/>)
 - Need a credit card and a phone number to register
2. Sign in to Oracle Cloud console using tenant id and credentials generated from step 1
(<https://www.oracle.com/cloud/sign-in.html>)



ORACLE[®] Cloud

Cloud Account Name ?

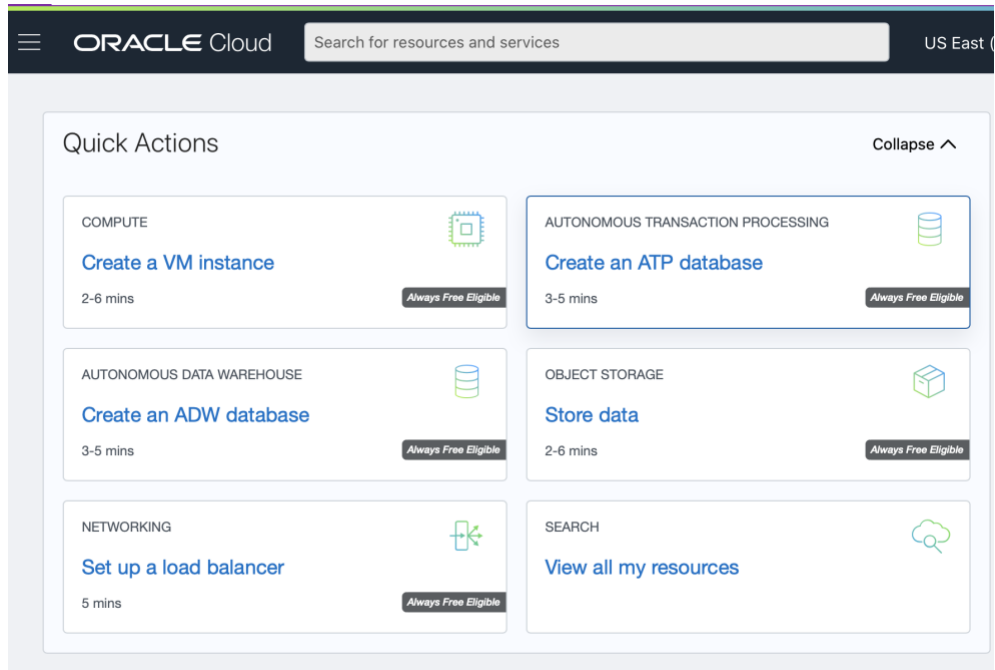
Next

[Forgot your cloud account name? Get help](#)

[Sign In using a Traditional Cloud Account](#)

3. Click on “Create an ATP database” to create an Autonomous Transaction Processing Database (Always Free).
Provide Admin Password (field can be found on the ATP form). Review the default values in the other fields on the form. You may keep the other details the same as suggested by Oracle.
Click “Create Autonomous Database”

Please refer to the following screenshots of the form.



Screenshots of the form:

Choose a workload type

Data Warehouse Configures the database for a decision support or data warehouse workload, with a bias towards large data scanning operations.	Transaction Processing Configures the database for a transactional workload, with a bias towards high volumes of random data access. ✓
---	--

Choose a deployment type

Shared Infrastructure Run Autonomous Database on shared Exadata infrastructure. ✓	Dedicated Infrastructure Run Autonomous Database on dedicated Exadata infrastructure.
---	---

Configure the database

Always Free ⓘ
☒ Show only Always Free configuration options

ⓘ If your Always Free Autonomous Database has no activity for 7 consecutive days, the database will be automatically stopped. Your data will be preserved, and you can restart the database to continue using it. If the database remains stopped for 3 months, it will be reclaimed. [Learn more](#)

Choose database version
 18c

OCPU count
 1
Always Free Autonomous Databases can utilize up to 1 core. The CPU core count cannot be adjusted.

Storage (TB)
 0.02
Always Free Autonomous Databases can utilize up to 0.02 TB (20 GB) of storage. The storage size cannot be adjusted.

☐ Auto scaling
Allows system to use up to three times the provisioned number of cores as the workload increases. [Learn more.](#)

Choose network access

☒ Allow secure access from everywhere
You can restrict access to specific IP addresses and VCNs.

☐ Virtual cloud network
Private access only, using a VCN.

☐ Configure access control rules ⓘ

Choose a license type

Bring Your Own License (BYOL)

Bring my organization's Oracle Database software licenses to the Database service. [Learn more.](#)


License Included

Subscribe to new Oracle Database software licenses and the Database service. ✓

4. The ATP status would change from Provisioning to Available:

ORACLE Cloud Search for resources and services

Autonomous Database » Autonomous Database Details



PROVISIONING...

DB 202005021917 **Always Free**

DB Connection Performance Hub Service Console S

Autonomous Database Information Tools Tags

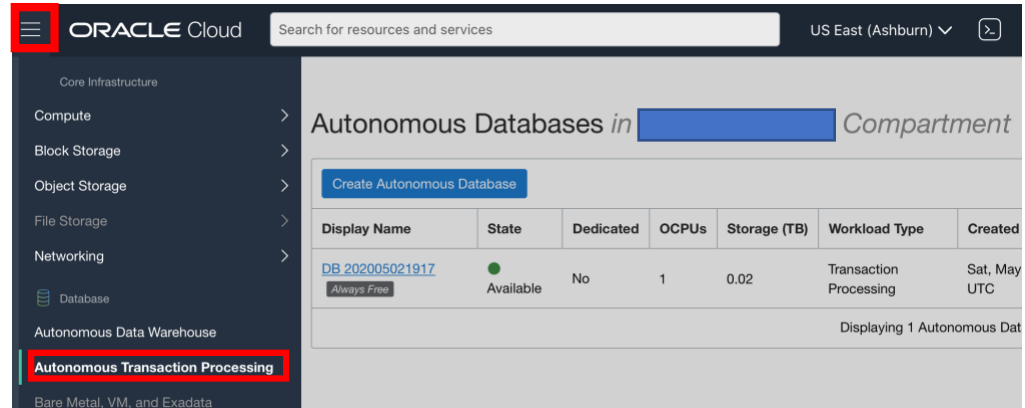
General Information

Database Name: DB202005021917

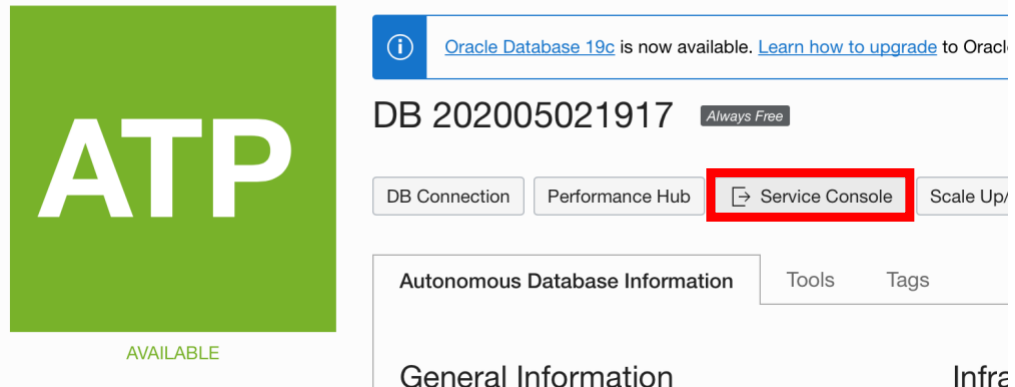
Workload Type: Transaction Processing

5. Create OML user: Now that we have our Autonomous Database available, the next step is to create a new Oracle Machine Learning (OML) user, since the ADMIN a/c is restricted from creating notebooks and jobs. The OML user can create notebooks and other jobs.

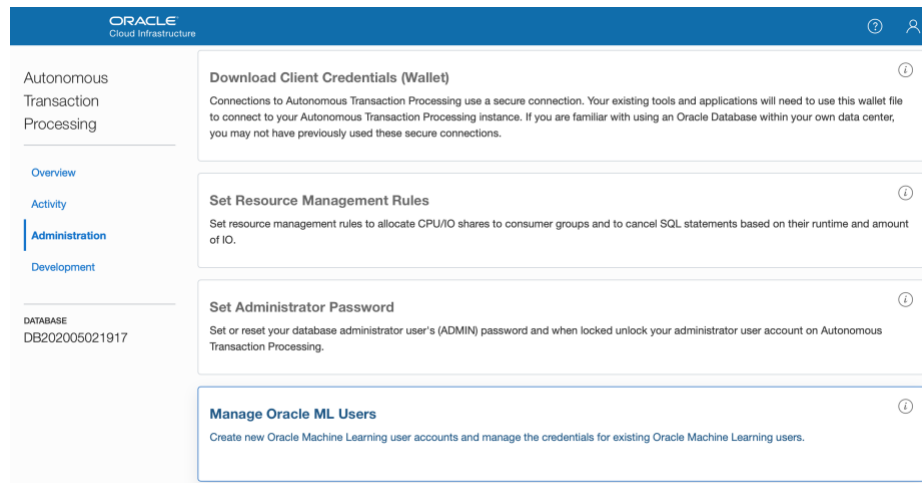
- a. Navigate to the new ATP and click on Service Console



Autonomous Database » Autonomous Database Details



- b. On the Service Console screen, click on Administration and then select Manage Oracle ML Users



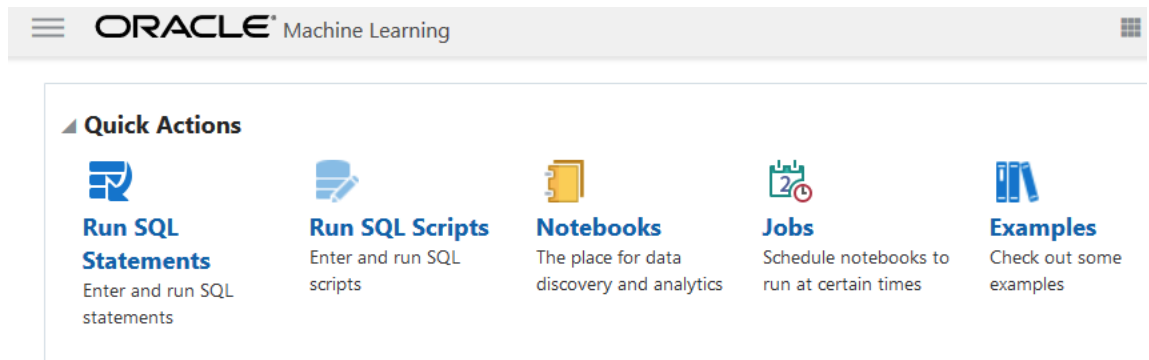
- c. Click on “+ Create” and create the required User by filling in details on the following screen.

The image shows two screenshots from the Oracle Machine Learning User Administration interface. The top screenshot is the 'Users' management page, which includes a '+ Create' button highlighted with a red box, along with 'Delete' and 'Show All Users' options. Below these are columns for 'User Name' and 'Full Name', with 'ADMIN' listed under 'User Name'. The bottom screenshot is the 'Create User' form, which contains input fields for 'Username', 'First Name', 'Last Name', and 'Email Address'. It also has a checkbox for 'Generate password and email account details to user' (which is checked), and fields for 'Password' and 'Confirm Password'.

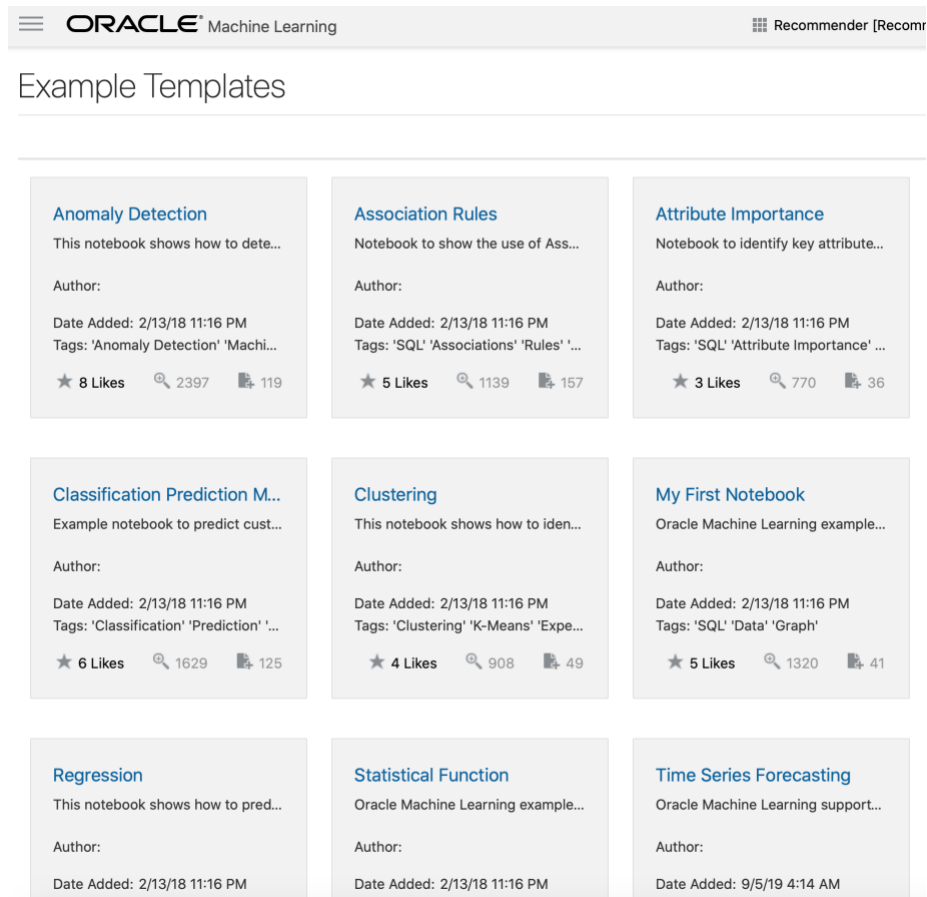
- d. Once we have the OML user working, we are all set to import data and use services such as SQL Developer Web and OML Notebooks that help implement machine learning algorithms through analytical PL/SQL functions.

Part II: Import data and implement ML algorithms/Recommendation Engine

1. Login using the URL and the OML user credentials provided by the OML ADMIN. The user will get an email containing the login URL and the password reset link.
2. Create Project and Workspaces
 - a. The OML service would create a default workspace and project, but you may create a new workspace and project.
For more details, refer to:
<https://www.oracle.com/webfolder/technetwork/tutorials/obe/cloud/oml/OMLPW-create-project-workspace/html/index.html>
3. Build tables and populate data using SQL Developer Web. You can also use OML notebooks for this step.
4. OML offers notebooks which we will use to build our recommendations. Notebooks are a fast and easy way to share analytical approaches and results of our machine learning pilot with our Data Management and Business Process Modeling audience
 - a. Create a notebook
 - b. Open the notebook in notebook editor and build your model
For more, refer to:
<https://www.oracle.com/webfolder/technetwork/tutorials/obe/cloud/oml/OMLCR-create-run-notebooks/html/index.html>



5. Oracle Machine Learning notebooks service provides example templates, under the “Examples” section, that may be used to implement some popular algorithms and machine learning techniques.



6. We want the sales team to have the ability to make data-informed recommendations to clients and increase cross-sell probability. To demo a pilot, we decided to use the Association Rules machine learning technique and build association rules models using the Apriori algorithm for the recommendations.

The algorithm requires data containing Case ID/Value pairs. A Case ID column such as customer id and a value column such as Product ID or Product name can be used.

For more details on the algorithm, refer to:

<https://docs.oracle.com/en/database/oracle/oracle-database/12.2/dmcon/apriori.html>

7. Join our TRANSACTIONS and PRODUCT table data and load into the SALES_TRANS_CUST table. The model uses this table as the single source of data from where it can build the Association Rules (AR) model.

```
CREATE OR REPLACE VIEW SALES_TRANS_CUST AS
  SELECT DISTINCT CUSTOMER_ID, PRODUCT_NAME, PRODUCT_LINE
    FROM (SELECT A.CUSTOMER_ID, B.PRODUCT_NAME,
  B.PRODUCT_LINE
    FROM transactions A, product B
   WHERE A.PRODUCT_ID = B.PRODUCT_ID
  );
```


8. Drop the current AR_SH_SAMPLE_SETTINGS table which contains the AR Model settings

```
-- Drop table AR_SH_SAMPLE_SETTINGS

BEGIN
    EXECUTE IMMEDIATE 'DROP Table AR_SH_SAMPLE_SETTINGS';
    EXCEPTION
        WHEN OTHERS THEN NULL;
END;
```

9. Build a new AR_SH_SAMPLE_SETTINGS table which will contain our AR model settings

```
-- Create table AR_SH_SAMPLE_SETTINGS

CREATE TABLE AR_SH_SAMPLE_SETTINGS (
    SETTING_NAME  VARCHAR2(30),
    SETTING_VALUE VARCHAR2(4000));
```

10. Insert Association Rules model settings data such as minimum support, minimum confidence, etc., into AR_SH_SAMPLE_SETTINGS. These parameters can be changed as required to limit build time or Association Rules

```
-- Insert data into AR_SH_SAMPLE_SETTINGS

BEGIN
    INSERT INTO AR_SH_SAMPLE_SETTINGS VALUES
    (DBMS_DATA_MINING.ASSO_MIN_SUPPORT,0.04);
    INSERT INTO AR_SH_SAMPLE_SETTINGS VALUES
    (DBMS_DATA_MINING.ASSO_MIN_CONFIDENCE,0.1);
    INSERT INTO AR_SH_SAMPLE_SETTINGS VALUES
    (DBMS_DATA_MINING.ASSO_MAX_RULE_LENGTH,2);
    INSERT INTO AR_SH_SAMPLE_SETTINGS VALUES
    (DBMS_DATA_MINING.ODMS_ITEM_ID_COLUMN_NAME,
    'PRODUCT_NAME');
    COMMIT;
END;
```

11. Now we can create our AR model using the Association mining function supported by Oracle Data Mining. We need to provide the model name, mining function, data table name containing the sales transactions, case id column name, and table name containing model settings for the AR model.

```
-- Create model for Market Basket Analysis using Sales
history transactional data

BEGIN
    DBMS_DATA_MINING.CREATE_MODEL (
        MODEL_NAME          => 'AR_SH_SAMPLE',
        MINING_FUNCTION      => DBMS_DATA_MINING.ASSOCIATION,
        DATA_TABLE_NAME    => 'SALES_TRANS_CUST',
        CASE_ID_COLUMN_NAME => 'CUSTOMER_ID',
        SETTINGS_TABLE_NAME => 'AR_SH_SAMPLE_SETTINGS'
    );
END;
```

12. Now that we have our model, we can use it to see what items to recommend when customer is interested in a particular product, say “Math Mayhem - Fun with Fractions” from our sales transaction data. In terms of the Apriori, the antecedent is “Fun with Fractions” and we want the consequent and the associated Lift. The AR Example provided by OML has a template which we will update with our antecedent.

```
-- What items should we recommend when customer is
interested in "Math Mayhem - Fun with Fractions"?
-- Assume the antecedent item is "Math Mayhem - Fun with
Fractions".
-- Since the number of items in antecedent is 1, the
number_of_items is 2.
-- choose top 5, ordered by rule lift to see top
recommendations by lift
-- Change sort to RULE_SUPPORT to see top recommendations
by support

SELECT ROWNUM RANK,
       CONSEQUENT_NAME RECOMMENDATION,
       NUMBER_OF_ITEMS NUM,
       ROUND(RULE_SUPPORT, 3) SUPPORT,
       ROUND(RULE_CONFIDENCE, 3) CONFIDENCE,
       ROUND(RULE_LIFT, 3) LIFT,
       ROUND(RULE_REVCNFIGIDENCE, 3) REVERSE_CONFIDENCE
FROM (SELECT * FROM DM$VRAR_SH_SAMPLE
      WHERE NUMBER_OF_ITEMS = 2
```

```

AND EXTRACT(antecedent, '//item[item_name="Math Mayhem
- Fun with Fractions"]') IS NOT NULL
ORDER BY RULE_LIFT DESC, NUMBER_OF_ITEMS)
WHERE ROWNUM <= 5;

```

Output:

RANK	RECOMMENDATION	NUM	SUPPORT	CONFIDENCE
LIFT	REVERSE_CONFIDENCE			
1	Chemistry Creations - Dinosaur Wonders	2	0.125	1
8				

The Transactions Data:

```
select * from SALES_TRANS_CUST;
```

CUSTOMER_ID	PRODUCT_NAME	PRODUCT_LINE
100003	Playful Pets - Benny the Beagle	Playful Pets - Dogs
100004	Block Bananza - Fall Fun - Fall 2019	Block Bananza - Seasons
100005	Block Bananza - Fall Fun - Fall 2019	Block Bananza - Seasons
100007	Playful Pets - Benny the Beagle	Playful Pets - Dogs
100002	Math Mayhem - Fun with Fractions	Math Mayhem - Basics
100002	Chemistry Creations - Dinosaur Wonders	Chemistry Creations - Animals
100006	Chemistry Creations - Cooking Class	Chemistry Creations - Cooking
100001	Block Bananza - Spring Sun - Spring 2020	Block Bananza - Seasons
100001	Dont Get Bored Board Game	Board Game - Family
100000	Block Bananza - Fall Fun - Fall 2019	Block Bananza - Seasons

A few things to note:

- An important constraint to note is that Oracle Data Mining supports association rules that have one or more items in the antecedent and a single item in the consequent.
- We could add more values to the antecedent to do a more complex market basket analysis, and we could mine more association rules. However, as the number of unique items increases to x , total number of itemsets increases to 2^x , and the algorithm may require significantly more build time. In such a scenario, we could reduce the itemsets considered by increasing the minimum support and minimum confidence and decreasing the max rule length parameters in the model settings table as well as by decreasing the number of items in our recommendation query to 2.
- Although, Oracle Cloud's ATP services are always free, some OML services are currently limited to first 30 days.
- For more details on the Apriori algorithm and the associated terms, please refer to: <https://docs.oracle.com/en/database/oracle/oracle-database/12.2/dmcon/apriori.html>

References:

1. <https://www.oracle.com/webfolder/technetwork/tutorials/obe/cloud/oml/OMLPW-create-project-workspace/html/index.html>
2. <https://www.oracle.com/webfolder/technetwork/tutorials/obe/cloud/oml/OMLCR-create-run-notebooks/html/index.html>
3. <https://docs.oracle.com/en/database/oracle/oracle-database/12.2/dmcon/apriori.html>
4. <https://docs.oracle.com/en/cloud/paas/autonomous-data-warehouse-cloud/omlug/using-oracle-machine-learning.pdf>
5. https://docs.oracle.com/en/database/oracle/oracle-database/12.2/arpls/DBMS_DATA_MINING.html