

Cluster Analysis App User Guide

Post installation instructions

Assign “ClusterPac Admin” permission set to administrators. This permission set allows users to create cluster models.

Assign “ClusterPac User” permission set to users. This permission set allows users to run models and view job results.

Overview

[Cluster analysis](#) is a task of grouping similar objects into a number of groups called clusters. Cluster Analysis Application will group your data into clusters using unsupervised machine learning algorithms, such as [K-Means](#) and [K-Medoids](#). To get more details on these algorithms and methodology refer to the “Cluster Analysis in Salesforce” document attached to the AppExchange listing.

Clustering of your data is performed in 3 steps:

1. Create cluster model

On this step you will choose the source object, select the fields which will be used in record similarity (distance) calculations, specify filter criteria and distance type for each field and choose the clustering algorithm. You can also specify fields you want to predict values for.

2. Run cluster model

Here you will specify clustering algorithm parameters, such as the number of clusters you want your data to group into. After this you will start the asynchronous job which will perform the calculations.

3. View and export results

After the calculations are complete you will see the formed clusters and cluster centers (centroids). Each source record will be mapped to a corresponding cluster and this information will be stored in the `clustan__ClusterJobResult__c` object. A visual representation of all clusters will be displayed using t-SNE technique.

Once the clustering of your data is completed you can calculate field value predictions or find similar records. Predictions and similar records are calculated using [k-nearest neighbors](#) algorithm.

Create cluster model

1. In the app launcher navigate to “Cluster Analysis” application. Go to “Cluster Models” tab and click the “New” button. Create model UI will appear:

Cluster Analysis

Cluster Models

Cluster Jobs

Cluster Model

Select data for clustering

Configure cluster model

* Select a cluster object:

Select an object

Cancel

→ Next

- In the dropdown select a Salesforce object which records you want to cluster. A list of fields, filter options and SOQL query editor will appear with pre-selected Id and Name fields:

Cluster Analysis

Cluster Models

Cluster Jobs

Cluster Model

Select data for clustering

Configure cluster model

* Select a cluster object:

Lead (Lead)

Field Search

Filter

+ Add Condition

* SOQL Query

SELECT Id, Name FROM Lead

Lead

Annual Revenue

AnnualRevenue

City

City

Clean Status

CleanStatus

Company

Company

Company D-U-N-S Number

CompanyDunsNumber

Converted Account ID

ConvertedAccountID

Cancel

→ Next

- From the list of fields choose fields which you want to use for clustering. You can use fields from the current and parent objects. SOQL query will change accordingly after you check or uncheck fields:

The screenshot shows the 'Cluster Model' configuration interface. At the top, there are tabs: 'Cluster Analysis', 'Cluster Models' (selected), and 'Cluster Jobs'. Below the tabs, there are two progress bars: 'Select data for clustering' (active) and 'Configure cluster model'. The main area is titled 'Cluster Model'. It includes a dropdown menu for 'Select a cluster object:' with 'Lead (Lead)' selected. Below this is a 'Field Search' input field. A list of fields is shown under the 'Lead' object, with 'Annual Revenue' checked. To the right, there is a 'Filter' section with an '+ Add Condition' button. Below the filter, there is a text area for the 'SOQL Query' containing the query: 'SELECT AnnualRevenue, Id, Industry, Name, NumberOfEmployees, Title FROM Lead'. At the bottom, there are 'Cancel' and 'Next' buttons.

- Optionally you can specify filter conditions. Click “Add Condition” button to create a new filter condition. Choose a field, condition operation and filter value:

The screenshot shows the 'Cluster Model' configuration interface, now in the 'Configure cluster model' step. The 'Select data for clustering' progress bar is still active. The 'Select a cluster object:' dropdown remains 'Lead (Lead)'. The 'Field Search' input field is empty. The list of fields under 'Lead' is still visible, with 'Annual Revenue' checked. The 'Filter' section now shows a condition: '1 Country Equals Canada'. There is a '+ Add Condition' button below the filter. The 'SOQL Query' text area now contains the query: 'SELECT AnnualRevenue, Id, Industry, Name, NumberOfEmployees, Title FROM Lead WHERE Country = 'Canada''. At the bottom, there are 'Cancel' and 'Next' buttons.

You can also edit SOQL query directly in the SOQL query editor. Please note that not all SOQL queries can be processed. Some SOQL functions such as aggregate (COUNT, SUM, etc) or FORMAT are not supported. A SOQL parsing error will be displayed in case if there is a problem with the query.

- Click “Next” to go to the model configuration screen
- On this screen enter the model name, clustering algorithm and the default number of clusters. “K-Medoids” algorithm can be used for most of the cases. You can use “K-Means” if your model has only numeric or long text clustering fields.

Cluster Analysis Cluster Models Cluster Jobs

Cluster Model

Configure cluster model

*Model name: Leads 4 fields model

*Clustering Algorithm: K-Medoids

*Number of clusters: 4

Model description

Field	Distance function or field type	Field value weight
Annual Revenue <i>Lead.AnnualRevenue</i>	Numeric	100%
Lead ID <i>Lead.Id</i>	None	
Industry <i>Lead.Industry</i>	Category	100%

Distance functions and field types

K-medoids uses Gower distance function to combine weighted values of different field types. K-means only works with Numeric fields and doesn't support weight values

Numeric: Manhattan or Euclidean distance function will be used to measure similarity of values

Text: Levenstein distance function will be

Cancel Back Save

- Distance functions were automatically populated based on the field type. You can change the distance function/field type using the dropdown besides each field. The following values are available:

Numeric	Manhattan or Euclidean distance function will be used to measure similarity of values
Text	Levenstein distance function will be used to measure similarity of values. Not supported in K-means
Category	An Equal (0)/Not equal (1) comparison will be used to measure similarity of values. Not supported in K-means
Long Text	TF-IDF algorithm (based on identical keywords) will be used to measure similarity.
None	The field will be skipped for similarity calculation. This is usually set for ID and name fields which are required to identify records
Cluster result output	A cluster number will be stored in this field during cluster calculations. The field name must start with 'ClusterNumber'. Warning: Any existing values in this field will be

	overwritten . Refer to “Populating cluster number into the source object” section for more details
Prediction	Cluster Analysis will try to predict the value of this field using K-Neareast Neighbors algorithm. Display only, the field value will not be updated or saved

Optionally you can specify weight for each field. To do this enter a decimal value in the 0..1 range. Default weight value in 1 which corresponds to 100%. This value will be used in distance calculations.

8. Click Save.

Run cluster model

There are 2 ways to run a model. First option is to run a model from the view model screen. To do this go to “Cluster Models” tab and click on a model which you want to run. You will be navigated to the view model screen. Click “Run model” button in the top right corner:

Cluster Analysis
Cluster Models
Cluster Jobs

Cluster Model
Leads 4 fields model

Run model
Edit
Delete

ClusterModel Name
Leads 4 fields model

Algorithm
K-Medoids

Description

Object
Lead

Default Number Of Clusters
4

Owner
User User

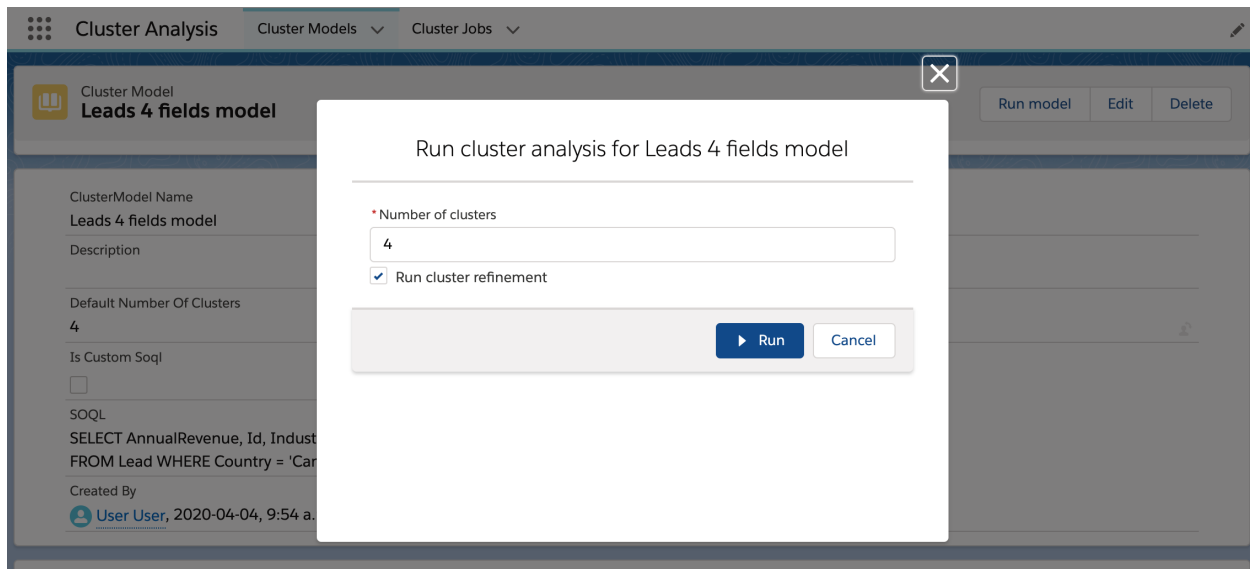
Is Custom Soql
☐

Last Modified By
User User, 2020-04-04, 9:54 a.m.

SOQL
SELECT AnnualRevenue, Id, Industry, Name, NumberOfEmployees, Title
FROM Lead WHERE Country = 'Canada'

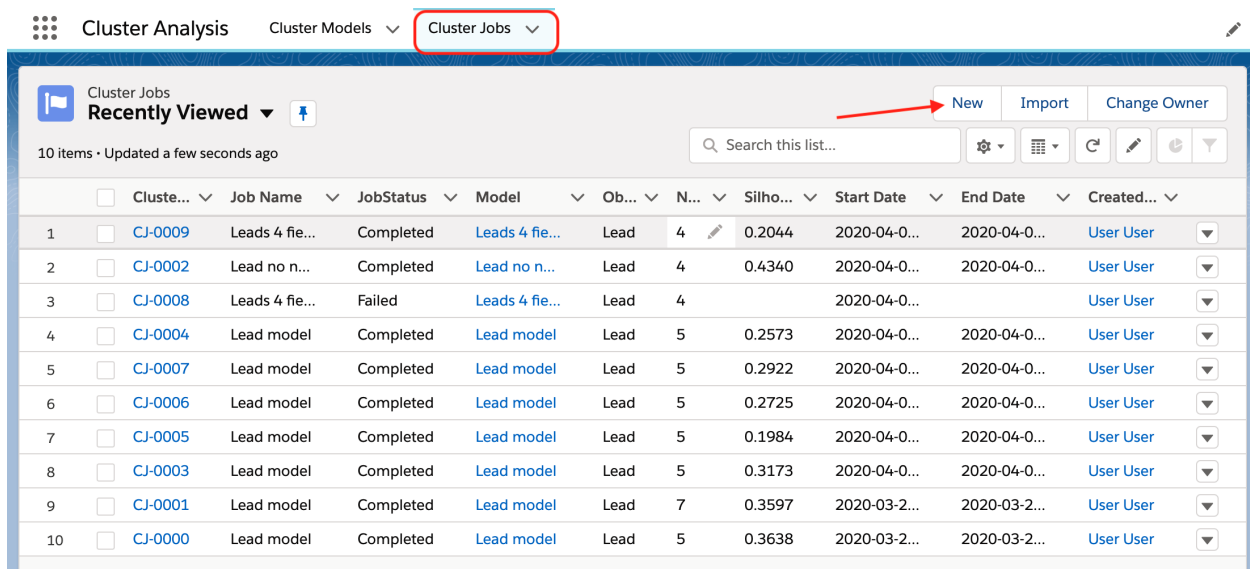
Created By
User User, 2020-04-04, 9:54 a.m.

A dialog box will appear where you can change clustering algorithm parameters:



You can change the number of clusters or specify whether or not to run refinement step. Click “Run” once you enter parameters.

Another option to run a model is to create a new cluster job. To do this go to “Cluster Jobs” tab and click “New” button:



You will be navigated to the new cluster job screen. On this screen choose a model from the list of models on the left, then enter algorithm parameters on the right and click “Run”:

Cluster Analysis Cluster Models Cluster Jobs

New cluster job

Cluster models: Select a model to run

Lead

Lead model

Lead no name model

Leads 4 fields model

Parameters: Enter algorithm parameters and click "Run"

* Number of clusters

4

☒ Run cluster refinement

Cancel Run

After you click "Run" a new cluster job will be created and you will be navigated to the job view screen:

Cluster Job CJ-0008 Delete

Job details:

Job Name

Leads 4 fields model

Job Number

CJ-0008

Job Status

New

Object Name

Lead

Model

Leads 4 fields model

Algorithm

K-Medoids

Number of clusters

4

Clusters:

Cluster job is running

Cancel job

The job is executed asynchronously by batch apex classes. The calculations might take some time depending on the number of records in the dataset, clustering algorithm and the number of clustering fields. The dataset of 1000 records will be processed in 5-10 minutes depending on the current instance load. The processing time of a 100K dataset varies from 40 minutes up to 2 hours. K-Medoids algorithm is usually faster than K-Means because it uses a subset of records in most of the calculations.

The clustering of LongText fields takes significantly more time than Numeric, Category or Text fields due to the complexity of TF-IDF calculations. The processing of 1000 records with long text data might take up to 1 hour with K-Medoids and up to several hours with K-Means. You can cancel the job using the "Cancel job" button. When calculations are finished the results will be displayed automatically.

View job details

You can see a list of “In Progress” or “Completed” jobs in the “Cluster Jobs” tab. Click on any job from the list to display the results. You will be navigated to the view job screen:

Cluster Analysis

Cluster Models

Cluster Jobs

Cluster Job
CJ-0009

Delete

Job details:

Job Name

Leads 4 fields model

Job Status

Completed

Model

Leads 4 fields model

Number of clusters

4

Number of records

1000

Start date

2020-04-04, 10:10 a.m.

Job Status Details

Cluster job finished

Job Number

CJ-0009

Object Name

Lead

Algorithm

K-Medoids

Silhouette score

0.2044

Created by

User User, 2020-04-04, 10:10 a.m.

End date

2020-04-04, 10:11 a.m.

Clusters:

Cluster 0

Record count: 271

Cluster centroid record:

Annual... 129000

Id: 00Q1100000D2ITJEAJ

Industry: Apparel

Name: Vittoria Darracott

Numb... 521

Title: Food Chemist

Cluster 1

Record count: 221

Cluster centroid record:

Annual... 177798

Id: 00Q1100000D2ISFEAZ

Industry: Agriculture

Name: Jarvis Ruddiman

Numb... 531

Title: Software Engineer III

Cluster 2

Record count: 259

Cluster centroid record:

Annual... 84723

Id: 00Q1100000D2IYdEAJ

Industry: Healthcare

Name: Allister Rozenbaum

Numb... 419

Title: VP Marketing

Cluster 3

Record count: 249

Cluster centroid record:

Annual... 37366

Id: 00Q1100000D2IMWEAZ

Industry: Technology

Name: Sandra Sommerton

Numb... 555

Title: Staff Scientist

If the job is completed you will see the calculation details such as silhouette score and the list of clusters with the record count and centroid records for each cluster.

By default the cluster job name will be the same as the cluster model name. You can rename the job using the pencil icon besides the job name.

t-SNE graph will be displayed below. t-SNE calculations might take a few minutes. The progress will be displayed on the graph itself - you will see the data points changing positions. Once the data points on the screen are stable the calculation is complete:



Please note that t-SNE reduces dimensionality of the data for visualization in a low-dimensional space, 2D in this case. t-SNE does not preserve distances, it only preserves nearest neighbors, so visualization of clusters might not match with the clustering algorithm results. t-SNE is a probabilistic algorithm, each consecutive run might produce a different output. You can change t-SNE graph parameters, such as: collide option, epsilon and perplexity. The graph will be recalculated accordingly. For more information on this technique and parameters refer to [t-SNE wiki page](#).

The graph will display 700 random data points (records) by default. You can change this in "Cluster Setting" custom metadata type in the setup, parameter name "TSNE Size". Please note that the larger value will require more memory and CPU usage by the browser which might slow down your system. The calculations are performed by the javascript in the browser. The maximum recommended value is 1500

Find similar records

You can access "Find Similar Records" from cluster model or job details page or using a lightning record page component

Cluster model or job details page

Open model or job details in “Cluster Analysis” app. Click “Find Similar Records” button on the highlights panel:

The screenshot shows the 'Cluster Analysis' app interface. At the top, there's a navigation bar with 'Cluster Analysis', 'Cluster Models', and 'Cluster Jobs'. Below this, a header bar displays 'Cluster Job CJ-0002' and a 'Find Similar Records' button, which is highlighted with a red circle and an arrow. The main content area is divided into two sections: 'Job details' and 'Clusters'.

Job details:

Job Name	Job Number
News	CJ-0002

Job Status: Completed

Model: News

Number of clusters: 20

Number of records: 4224

Start date: 12/18/2020, 10:10 AM

Job Status Details: Cluster job finished

Object Name: clustan__news__c

Algorithm: K-Medoids

Silhouette score: 0.0611

Created by: User User, 12/18/2020, 10:10 AM

End date: 12/18/2020, 5:06 PM

Clusters:

Cluster	Record count	Cluster centroid record:
Cluster 0	240	Record ID: a0754000002qTV6AAM
Cluster 1	82	Record ID: a0754000002qT6vAAE

Cluster 0 details:

title: A closer look at SNC-Lavalin's sometimes murky past

description: One of Canada's biggest engineering companies is at the centre of what appears to be a growing scandal engulfing Prime Minister Justin Trudeau and his government.

text: Top 20 stemmed keywords: snc-lavalin, libya, charg, contract, libyan, employe, bangladesh, morin, corrupt, plead, compani,

Cluster 1 details:

title: Iran sending black boxes of downed plane to Ukraine, report says

description: Iran is sending to Ukraine the black boxes of the Ukrainian passenger plane that its military accidentally shot down this month, Tasnim news agency reported on Saturday.

text: Top 20 stemmed keywords: iran, tasnim, box, plane, black, franc, ps752, flight, rezaifar, accident, kyiv, ukrainian,

In the popup dialog enter the record name or id, choose the record of interest in the search results and click “Find” button:

Find Similar Records

* Search news

a0754000002qTsdAAE

Q

* Number of records

15

Find

Young scientists pen letter to Trudeau urging environmental assessment overhaul

clustan__news__c

Cancel

Similar records will be retrieved and displayed below sorted by similarity percentage:

Find Similar Records

* Search news * Number of records

Young scientists pen letter to Trudeau urging environmental assessment overhaul ✕

Find

Young scientists pen letter to Trudeau urging environmental assessment overhaul Cluster: 16 ■ Record ID: a0754000002qTsiAAE title: Young scientists pen letter to Trudeau urging environmental assessment overhaul description: More than 1,300 early-career scientists have signed an open letter to Justin Trudeau and several of his ministers saying environmental assessments need a "higher bar." text: Top 20 stemmed keywords: scientif, jacob, scientist, letter, lng, environment, assess, evid, northwest, pacif, long-form, ministri, research, project, incorpor, transpar, gatewai, censu, decision-mak, input Similarity: 99.24%	What you need to know about the Coastal GasLink pipeline conflict Cluster: 16 ■ Record ID: a0754000002qTlqAAM title: What you need to know about the Coastal GasLink pipeline conflict description: The conflict over a natural gas pipeline in northwestern British Columbia is the latest flashpoint between resource development and Indigenous rights and title in a province where large swaths of territory are not covered by any treaty. text: Top 20 stemmed keywords: wet'suwet'en, gaslink, hereditari, coastal, jan, chief, road, gidimt'en, rcmp, pipelin, checkpoint, tree, injunct, arrest, project, lng, territori, forest, horgan, camp Similarity: 9.41%
Trudeau's 'weak leadership' drove Teck mine away, Andrew Scheer says Cluster: 16 ■ Record ID: a0754000002qTL8AAM title: Trudeau's 'weak leadership' drove Teck mine away, Andrew Scheer says description: Outgoing Conservative Leader Andrew Scheer	Canada better on human rights, but not Indigenous rights: Amnesty International Cluster: 16 ■ Record ID: a0754000002qTnjAAE title: Canada better on human rights, but not Indigenous rights: Amnesty International description: Canada has made progress improving its human

Cancel

Lightning Record Page Component

To access “Find Similar Records” using the lightning record page component you will need to create and activate a lightning record page for your cluster object if you don’t have one. [This trail](#) describes how to do this. Once you created a lightning record page open it in design mode, find “clusterNearestNeighbors” custom managed component and drag it to the desired location. Enter cluster model name or job id of the completed cluster job. Optionally change the header label or number of similar records you want this component to display. Click “Save”:

Lightning App Builder | Pages | Lead

Desktop | Shrink To View

Analyze | Activation... | Save

Components

Search...

- ClusterJobDetails
- ClusterTSNE
- Image Carousel
- Custom - Managed (12)
 - Candidate Patient List View for Health Cloud
 - Care Plan Selector for Health Cloud
 - ClusterJobDetails
 - clusterNearestNeighbors**
 - clusterPredictResult
 - Patient Card for Health Cloud
 - Patient Detail for Health Cloud
 - Patient Info for Health Cloud
 - Patient List View for Health Cloud
 - Patient Navigation for Health Cloud
 - Patient Timeline for Health Cloud
 - Today View for Health Cloud

Lead: Calley Laffoley-Lane

Title: Geologist III | Company: Tagfeed | Phone: | Email: claffoleylane0@globo.com

Lead Owner: Iskander Mukhamedgaliyev | Lead Status: Open

Name: Calley Laffoley-Lane | Phone: | Email: claffoleylane0@globo.com | Lead Quality: \$\$\$\$

Address Information

Address: 8 Sunbrook Crossing, Roanoke, VA 24040 | Website: | Map: GAINSBORO, Roanoke

Additional Information

No. of Employees: 472 | Lead Source: | Annual Revenue: \$65,308 | Industry: Agriculture | Description: Hi since recruiter lead permission approve requisitions makes please make thanks recruiter

Lead Scoring

Decision Maker: | Decision Maker Score: 0.00 | Project Defined: | Project Defined Score: 0.00 | Not Budget: | Budget Score: 0.00

Page: clusterNearestNeighbors

* Cluster Job Id or Model Name: LeadsPrediction (This field is required.)

* Number of similar records: 10

* Header label: Similar records

Set Component Visibility

Filters: + Add Filter

If the page is activated activated similar records will be displayed on the record detail page:

Lead: Calley Laffoley-Lane

Title: Geologist III | Company: Tagfeed | Phone: | Email: claffoleylane0@globo.com

+ Follow | Convert | Edit | Delete

Similar records

Will Lanceter

Cluster: 0

Annual: 79486

Revenue: | Lead ID: 00Q5B000007AWHcUA0 | Industry: Banking | Full Name: Will Lanceter | Employees: 523

Similarity: 93.26%

Marja Avarne

Cluster: 3

Annual: 67658

Revenue: | Lead ID: 00Q5B000007AWCHUA4 | Industry: Other | Full Name: Marja Avarne | Employees: 334

Similarity: 92.37%

Lead Owner: Iskander Mukhamedgaliyev | Lead Status: Open

Name: Calley Laffoley-Lane | Phone: | Email: claffoleylane0@globo.com | Lead Quality: \$\$\$\$

Address Information

Address: 8 Sunbrook Crossing, Roanoke, VA 24040 | Website: | Map: GAINSBORO, Roanoke

Predictions

You can access prediction functionality from cluster model or job details page, using a lightning record page component or using a lightning quick action

Cluster model or job details page

Open model or job details in “Cluster Analysis” app. Click “Predict” button on the highlights panel:

The screenshot displays the 'Cluster Analysis' app interface. At the top, a header bar shows 'Cluster Job CJ-0000' on the left and three buttons ('Find Similar Records', 'Predict', 'Delete') on the right. The 'Predict' button is highlighted with a red circle and a red arrow pointing to it from below. The main content area is divided into two sections: 'Job details:' on the left and 'Clusters:' on the right.

Job details:

Job Name	LeadsPrediction
Job Status	Completed
Model	LeadsPrediction
Number of clusters	4
Number of records	1000
Start date	1/2/2021 5:37 AM
Job Status Details	Cluster job finished
Job Number	CJ-0000
Object Name	Lead
Algorithm	K-Medoids
Silhouette score	0.404
Created by	Iskander Mukhamedgaliyev, 1/2/2021 5:37 AM
End date	1/2/2021 5:39 AM

Clusters:

Cluster	Record count	Cluster centroid record:
Cluster 0	309	Annual: 67085 Revenue: 00Q5B000007AWL9UAO Lead ID: Technology Industry: Tobiah Warstall Full Name: 708 Employee es:
Cluster 1	265	Annual: 178002 Revenue: 00Q5B000007AWGtUAO Lead ID: Apparel Industry: Fleming Phizakarley Full Name: 352 Employee es:
Cluster 2	229	Annual: 159800 Revenue: 00Q5B000007AWDhUAO Lead ID: Agriculture Industry: Ailis Grayson Full Name: 831 Employee es:
Cluster 3	197	Annual: 70029 Revenue: 00Q5B000007AWH3UAO Lead ID: Banking Industry: Rennie Trimme Full Name: 49 Employee es:

In the popup dialog enter the record name or id, choose the record of interest in the search results and click “Predict” button:

Predict Record Cluster

* Search Lead

te

Temp Shapcote

Lead

Hans Texton

Lead

Jordon Tewelson

Lead

Remington Tedstone

Lead

Tedi Westphal

Lead

Cancel

Predict

Predictions will be calculated and displayed in a few moments:

Predict Record Cluster

Record name:Temp Shapcote

Model:LeadsPrediction

Predicted cluster:Cluster 0

Field predictions:

Industry

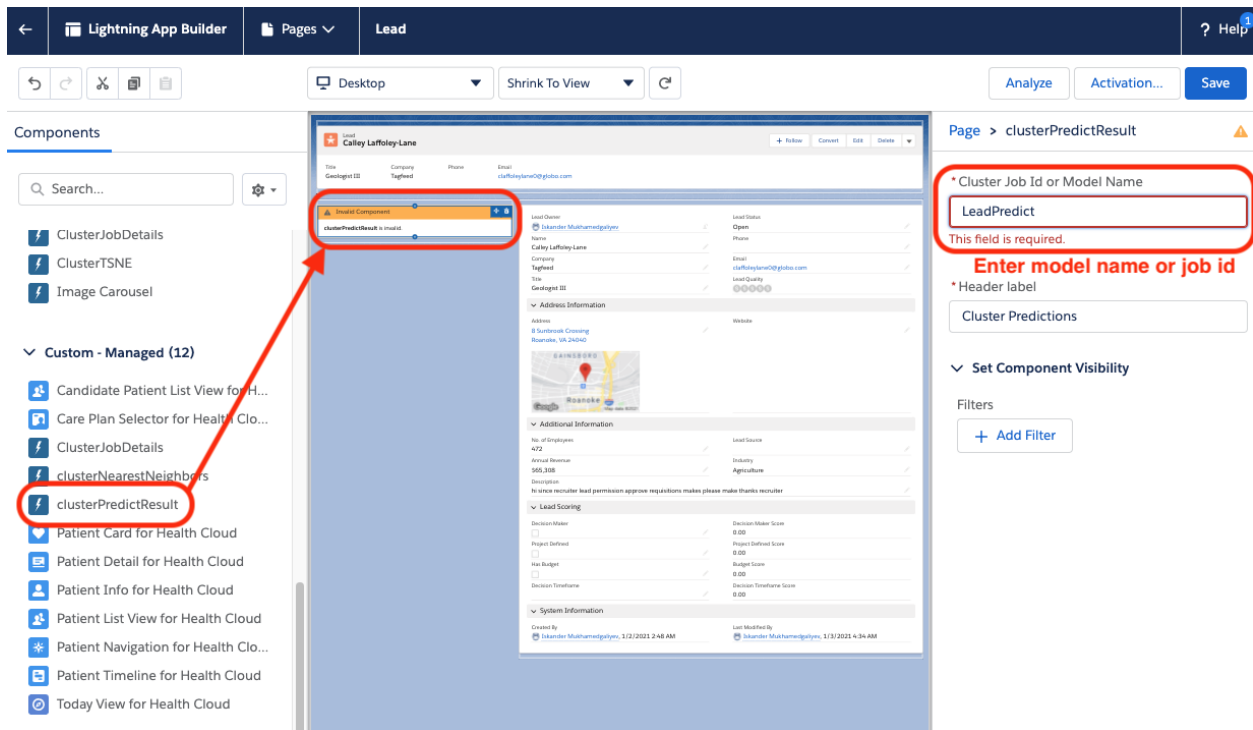
Predicted Value	Probability
Technology	40.70%
Other	13.24%
Banking	13.01%

Close

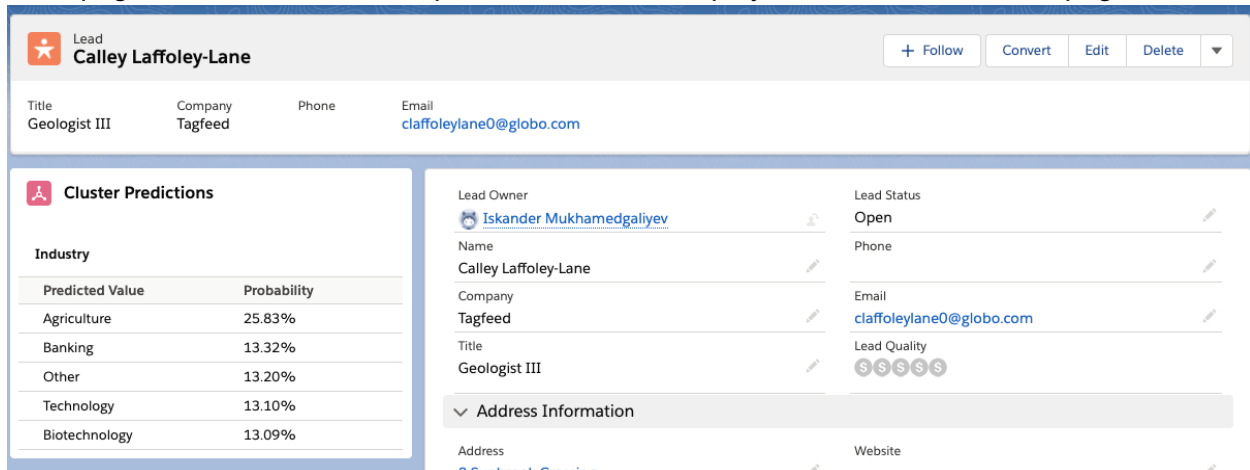
Predicted cluster will be displayed by default. If your model has prediction fields configured the predicted values and probabilities for each value will also be displayed.

Lightning Record Page Component

To access prediction functionality using the lightning record page component you will need to create and activate a lightning record page for your cluster object if you don't have one. [This trail](#) describes how to do this. Once you created a lightning record page open it in design mode, find "clusterPredictResult" custom managed component and drag it to the desired location. Enter cluster model name or job id of the completed cluster job. Optionally change the header label. Click "Save":



If the page is activated activated predictions will be displayed on the record detail page:



Lightning quick action

To access predictions using a lightning quick action you need to create prediction quick action on the cluster object first. To do this go to "Setup\Object Manager", find your cluster object and go to "Buttons, Links, and Actions". Click "New Action" and create the following quick action:

Action type: Lightning Component

Lightning Component: clustan:ClusterPredictQuickAction

Height: 450px

Label: Predict

Name (auto populated): Predict

Click Save:

The 'Enter Action Information' dialog box is shown with the following fields:

- Object Name:** Lead
- Action Type:** Lightning Component
- Lightning Component:** clustan:ClusterPredictQuickAction
- Height:** 450px
- Standard Label Type:** --None--
- Label:** Predict
- Name:** Predict
- Description:** (empty)
- Icon:** Change Icon

There are 'Save' and 'Cancel' buttons at the top and bottom of the dialog.

Next you will need to add this quick action to the page layout. Open the corresponding page layout in design mode and add this quick action:

The screenshot shows the 'Lead' object page in Salesforce Object Manager. The left sidebar lists various configuration areas, with 'Page Layouts' selected. The main content area shows the 'Mobile & Lightning Actions' section, which is currently in design mode. A red circle highlights the 'Predict' button in the 'Quick Actions in the Salesforce Classic' section. A red arrow points from this button to the 'Predict' button in the 'Salesforce Mobile and Lightning Experience Actions' section, indicating the process of adding the quick action to the mobile layout. The 'Predict' button is also highlighted in the 'Salesforce Mobile and Lightning Experience Actions' section.

Click Save. Predict button should appear on the record details page:

Lead

Calley Laffoley-Lane

+ Follow

Predict

Change Owner

Delete

Title

Geologist III

Company

Tagfeed

Phone

Email

claffoleylane0@globo.com

Cluster Predictions

Industry

Predicted Value	Probability
Agriculture	25.83%
Banking	13.32%
Other	13.20%
Technology	13.10%
Biotechnology	13.09%

Lead Owner

Iskander Mukhamedgaliyev

Name

Calley Laffoley-Lane

Company

Tagfeed

Title

Geologist III

Lead Status

Open

Phone

Email

claffoleylane0@globo.com

Lead Quality

\$\$\$

Address Information

Address

Website

When you click “Predict” a popup dialog will appear with the list of cluster models for your object. If you have only one model the list will have just one item. Select the model which you want to use for predictions and click “Predict” button:

Predict Record Cluster

Select model

Lead Industry&NumberOfEmployees Example Model

LeadsPrediction

Cancel

Predict

Prediction results will be displayed in a few moments:

Predict Record Cluster

Record name: Calley Laffoley-Lane

Model: LeadsPrediction

Predicted cluster: Cluster 3

Field predictions:

Industry

Predicted Value	Probability
Agriculture	25.83%
Banking	13.32%
Other	13.20%

Close

Predictions and similar records are calculated using [k-nearest neighbors](#) algorithm

Cluster job results

Clustering results will be stored in clustan__ClusterJobResult__c object. The object contains associations between the source object records and calculated clusters. This object has the following fields:

clustan__ClusterJob__c	Contains the id of the cluster job record which was created when you ran the model. Reference to clustan__ClusterJob__c
clustan__ClusterNumber__c	Contains the number of the record's cluster. Starts with 0
clustan__Cluster__c	Contains the id of the record's cluster. Reference to clustan__ClusterJobCluster__c
clustan__RecordId__c	Contains the record id of the cluster model object record. If Lead object was specified in the cluster model this field will contain a lead record id

clustan__RecordName__c	Contains the record name
clustan__Json__c	Contains the record data in json format. The data will only contain fields specified in the cluster model

Cluster job results will be displayed on the job details page. You can also export cluster job results to a csv file using the Salesforce report builder or the Data Loader tool. You need to filter the data on the job id field (clustan__ClusterJob__c) to retrieve results for a specific job (most cases).

Populating cluster numbers into the source object

It is possible to populate a cluster number value into the source (model) object records. To do this you will need to create a custom field in the source object which starts with "ClusterNumber" and add this field to the model with "Cluster result output" field type. Here is how to populate the cluster number into Lead records:

1. In the setup go to "Object Manager", click on "Lead" object and go to "Fields & Relationships"
2. Click "New" to create a new custom field.
3. Select "Number" data type and click "Next"
4. Enter "ClusterNumber" into the field label. Make sure that field name is also set to "ClusterNumber". You can change the field label if you want but the field name must start with "ClusterNumber":

Step 2. Enter the details Step 2 of 4

Previous **Next** Cancel

Field Label Cluster Number

Please enter the length of the number and the number of decimal places. For example, a number with a length of 8 and 2 decimal places can accept values up to "12345678.90".

Length 18 **Decimal Places** 0

Number of digits to the left of the decimal point Number of digits to the right of the decimal point

Field Name ClusterNumber

Description

Help Text

☐ **Required** Always require a value in this field in order to save a record

☐ **Unique** Do not allow duplicate values

☐ **External ID** Set this field as the unique record identifier from an external system

Default Value [Show Formula Editor](#)

Use formula syntax: Enclose text and picklist value API names in double quotes : ("the text"). include numbers without quotes

5. Click "Next"

- Set field level security for profiles and add this field to your page layout(s).
- Click “Save”. The field is created.
- Go to “Cluster Models” tab and edit an existing Lead model. Create a new Lead model if you don’t have one
- In the edit model screen find and select the “ClusterNumber__c” field. Click “Next”

The screenshot shows the 'Cluster Model' configuration interface. At the top, there are tabs for 'Cluster Analysis', 'Cluster Models', and 'Cluster Jobs'. The 'Cluster Models' tab is active. Below the tabs, there are two progress bars: 'Select data for clustering' (active) and 'Configure cluster model'. The 'Select data for clustering' section includes a dropdown for 'Select a cluster object' set to 'Lead (Lead)'. Below this is a 'Field Search' box and a list of fields under the 'Lead' object. The 'Cluster Number' field, labeled 'ClusterNumber__c', is selected and highlighted with a red box. To the right of the field list is a 'Filter' section with an 'Add Condition' button and a 'SOQL Query' box containing the query: 'SELECT ClusterNumber__c, Id, Industry, Name, NumberOfEmployees FROM Lead'. At the bottom right, there are 'Cancel' and 'Next' buttons, with a red arrow pointing to the 'Next' button.

- Specify “Cluster Result Output” for “ClusterNumber__c” field and click “Save”

The screenshot shows the 'Cluster Model' configuration interface, now in the 'Configure cluster model' step. The progress bar for 'Select data for clustering' is green with a checkmark, and the 'Configure cluster model' bar is blue. The 'Model name' is 'Lead Industry&NumberOfEmployees Example Model', the 'Clustering Algorithm' is 'K-Medoids', and the 'Number of clusters' is '5'. The 'Model description' is 'This model is created as an example'. Below this is a table with columns: 'Field', 'Distance function or field type', and 'Field value weight'. The 'Cluster Number' field, labeled 'Lead.ClusterNumber__c', is selected, and its 'Distance function or field type' is set to 'Cluster Result Output', which is highlighted with a red box. The 'Field value weight' is set to '100%'. At the bottom right, there are 'Cancel', 'Back', and 'Save' buttons, with a red arrow pointing to the 'Save' button. To the right of the table is a section titled 'Distance functions and field types' with text explaining the K-medoids algorithm and supported field types.

Field	Distance function or field type	Field value weight
Cluster Number <i>Lead.ClusterNumber__c</i>	Cluster Result Output	100%
Lead ID <i>Lead.Id</i>	Numeric	
Industry <i>Lead.Industry</i>	Category	
Full Name <i>Lead.Name</i>	Text	
	None	
	Cluster Result Output	

11. Click “Run model” to run the model and populate the cluster number into Lead.ClusterNumber__c field. Once the clustering job is finished the cluster number will be populated into ClusterNumber__c field for all processed Lead records:

The screenshot shows a lead record details page with tabs for Activity, Details, Chatter, and News. The Details tab is active. The record information is as follows:

Field	Value
Lead Owner	Iskander Mukhamedgaliyev
Name	Mr. Test Lead
Company	Test company
Title	VP Consulting
Cluster Number	4
Lead Status	Open
Phone	(829) 123-456
Email	test@example.com
Lead Quality	5 green circles, 1 grey circle

The 'Cluster Number' field is highlighted with a red rectangle. Below the record details is a section for 'Address Information' which is currently collapsed.

Please note that each time you run the model the previous values in this field will be overwritten.

Delete cluster job or model

To delete a cluster job go to the cluster job details page and click the “Delete” button. You will be redirected to a confirmation page. After confirming your action an asynchronous batch job will start which will delete this cluster job. Warning: **all associated clusters and job results will also be deleted.**

To delete a model go to the model details page and click the “Delete” button. After confirming your action an asynchronous batch job will start which will remove this cluster model. Warning: **all associated cluster jobs, clusters and job results will also be deleted.**

The delete operation in both cases might take several minutes depending on the size of the objects and number of records being deleted.

Appendix

Cluster objects ER Diagram

