MongoDB.

>

Products ~

Docs Home → Launch & Manage MongoDB → MongoDB Atlas

Perform Semantic Search with Atlas Vector Search

Overview

You can perform semantic search on data in your Atlas cluster running MongoDB v6.0.11, v7.0.2, or later using Atlas Vector Search. You can store vector embeddings for any kind of data along with other data in your collection on the Atlas cluster. Atlas Vector Search supports embeddings that are less than and equal to 2048 dimensions in width.

When you define an Atlas Vector Search index on your collection, you can seamlessly index vector data along with your other data and then perform semantic search against the indexed fields.

Atlas Vector Search uses the Hierarchical Navigable Small Worlds algorithm to perform the semantic search. You can use Atlas Vector Search support for ANN queries to search for results similar to a selected product, search for images, etc.

About Atlas Vector Search Indexes

You must index the fields that contain vector embeddings of <u>BSON</u> double data type for performing semantic search against the vector data. To perform a vector search against your data, create an Atlas Vector Search vector Search type index for your data. In the Atlas Vector Search index definition, you must index the field that contains the vector data as the vector type.

You can also optionally index fields for pre-filtering the data against which you want to run your Atlas Vector Search query. You can pre-filter your data by boolean, numeric, and string values. To pre-filter the data, you must index your boolean, numeric, and string fields as the filter type in the Atlas Vector Search index definition that indexes your vector data field.

To learn more about indexing your fields for Atlas Vector Search, see How to Index Fields for Vector Search.

About Atlas Vector Search Queries

Atlas Vector Search queries take the shape of an aggregation pipeline stage. The Atlas Vector Search stage must be the first stage in the pipeline. You can run Atlas Vector Search queries only against fields indexed as the vector type in a vector Search type index.

You can pre-filter data against which you want to perform semantic search using an MQL match expression that compares an indexed field with boolean, number, or string values. Atlas Vector Search supports some comparison query and aggregation pipeline operators in the filter in your \$vector Search query.

To learn more about pre-filtering and querying your data using Atlas Vector Search, see Run Vector Search Queries.

Procedure

This section describes how to index vector embeddings in your data on an Atlas cluster and run queries that search vector embeddings.

Define your Atlas Vector Search index.

In your Atlas Vector Search vector Search type index definition, you can index the field that contains the vector embeddings and the fields that you want to pre-filter your data by.

- a. Required. Index the field that contains vector data in your collection as the vector type (lines 3 - 8 below).
- b. Optional. Index boolean, numeric, and string fields as the filter type (lines 9 12 below) to enable vector search on filtered data.

The following is the syntax for an Atlas Vector Search index.

```
1
                                                                               ťa
    {
       "fields":[
3
           "type": "vector",
5
           "path": "<field-name>".
6
           "numDimensions": <number-of-dimensions>,
           "similarity": "euclidean | cosine | dotProduct"
8
9
           "type": "filter",
10
           "path": "<field-name>"
11
12
        },
13
```

On this page

Overview

Procedure

```
14 ]
15 }
```

To learn more, see How to Index Fields for Vector Search.

2 Construct your vector search query.

Use the \$vectorSearch pipeline stage in your query.

```
1
                                                                             ťa
2
       "$vectorSearch": {
         "index": "<index-name>",
3
4
         "path": "<field-to-search>",
         "queryVector": [<array-of-numbers>],
5
6
         "numCandidates": <number-of-candidates>,
         "limit": <number-of-results>,
7
         "filter": {<filter-specification>}
8
9
10 }
```

To learn more about this pipeline stage, see Run Vector Search Queries.

3 (Optional) Specify whether you want the score for the documents in the results.

Use vectorSearchScore in the \$project stage after your \$vectorSearch stage to retrieve the score for the documents in the results. To learn more, see Score the Documents in the Results.

4 Run your query.

Verify your query syntax and then run it using mongosh or a supported driver.

← Atlas Search Changelog

Create and Manage Atlas Vector Search Indexes

