# **CAS Registry Number Validator**

## **Introduction**

**Package Name: casrnvalidator**

**Description: casrnvalidator** is a Python package designed to facilitate working with Chemical Abstracts Service (CAS) Registry Numbers. It comprises two main modules, Validator and Search, each serving distinct functions related to CAS Registry Numbers.

### **1. Validator Module:**

The Validator module enables users to validate whether a given string corresponds to a valid CAS Registry Number format.

**Functions:**

**1. Function: validateCAS\_RN (cas\_number: input\_str)**

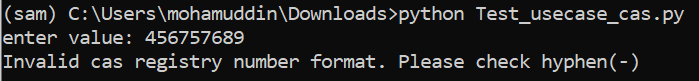
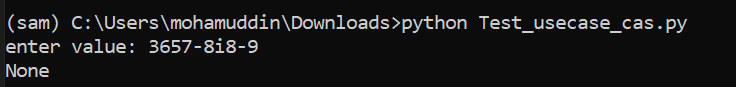
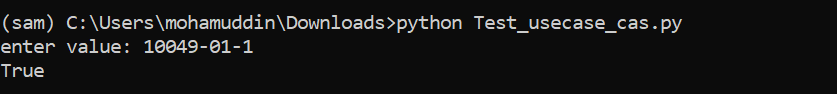
**Description:** Validates whether the input string cas\_number conforms to the CAS Registry Number format.

**input\_str**: The CAS Registry Number.

**Returns Type:** bool

**Returns:** Returns True if the input is a valid CAS Registry Number, otherwise False.

**Possible outputs:**



**2. Function: validateCASwithMessage (cas\_number: input\_str)**

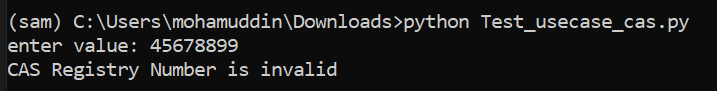
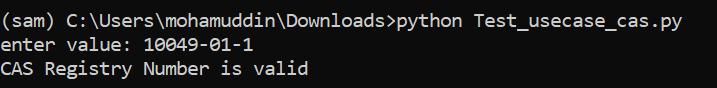
**Description:** Validates whether the input string cas\_number confirms to the CAS Registry Number format.

**input\_str**: The CAS Registry Number.

**Returns Type:** String

**Returns:** Return a valid or invalid message if the input is a valid CAS Registry Number.

**Possible outputs:**



### **2. Search Module:**

The Search module provides multiple functions. These functions provide a comprehensive set of tools for searching and retrieving information related to chemical compounds based on CAS Registry Numbers, synonyms, and chemical formulas. It utilizes the dataset to fetch detailed information about the compounds.

### **Functions:**

**1. Function: getChemicalFormulaWithCAS(input\_str)**

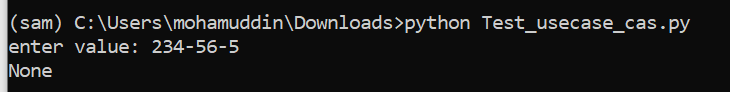
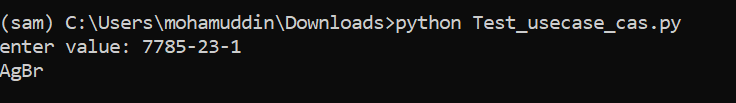
**Description:** Retrieves the chemical formula associated with the provided CAS Registry Number

**input\_str**: The CAS Registry Number.

**Returns Type:** String

**Returns:** The chemical formula corresponding to the given CAS Registry Number

**Possible outputs:**

.

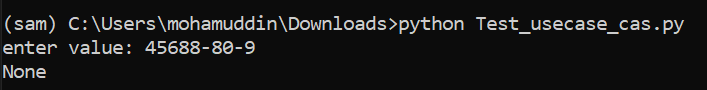
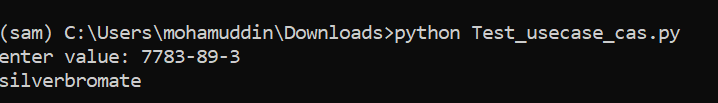
**2. Function: getSynonymsWithCAS(input\_str)**

**Description:** Fetches synonyms of the compound identified by the provided CAS Registry Number.**input\_str**: The CAS Registry Number.

**Returns Type:** String

**Returns:** Synonyms of the compound associated with the given CAS Registry Number

**Possible outputs:**

.

### **3. Function: getChemicalFormulaWithSynonyms(input\_str)**

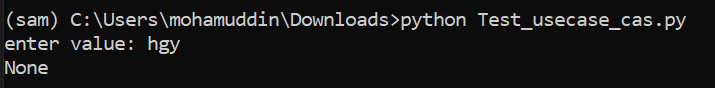
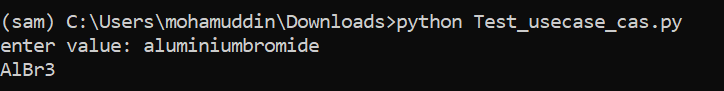
**Description:** Retrieves the chemical formula based on the synonyms provided.

**input\_str**: The synonym or list of synonyms.

**Returns Type:** String

**Returns:** The chemical formula corresponding to the provided synonym(s)

**Possible outputs:**

.

### **4. Function: getCASNumberWithSynonyms(input\_str)**

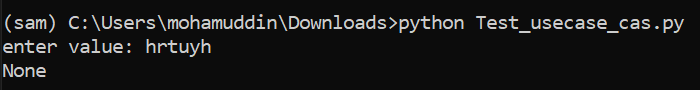
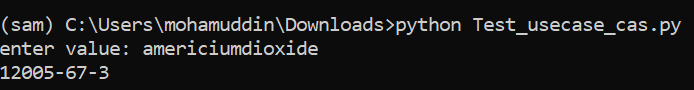
**Description:** Retrieves the CAS Registry Number associated with the given synonym(s).

**input\_str**: The synonym or list of synonyms.

**Returns Type:** String

**Returns:** The CAS Registry Number(s) corresponding to the provided synonym(s)

**Possible outputs:**

.

### **5. Function: getCASNumberWithChemicalFormula(input\_str)**

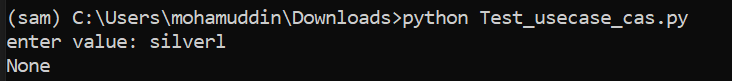
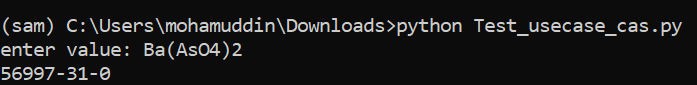
**Description:** Retrieves the CAS Registry Number associated with the provided chemical formula

**input\_str**: The chemical formula.

**Returns Type:** String

**Returns:** The CAS Registry Number corresponding to the provided chemical formula

**Possible outputs:**

.

### **6. Function: getSynonymsWithChemicalFormula(input\_str)**

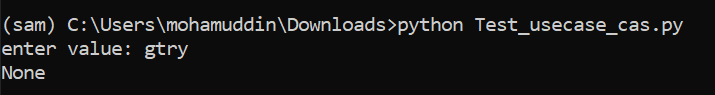
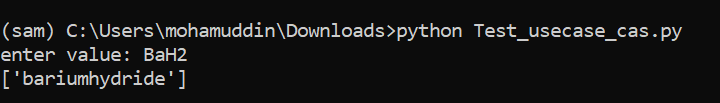
**Description:** Retrieves synonyms based on the provided chemical formula.

**input\_str**: The chemical formula.

**Returns Type:** String

**Returns:** Synonyms associated with the provided chemical formula

**Possible outputs:**

.

### **7. Function: search\_by\_cas\_number(input\_str)**

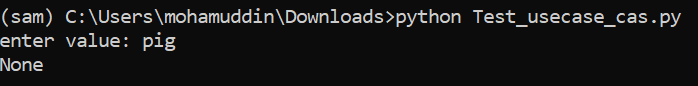
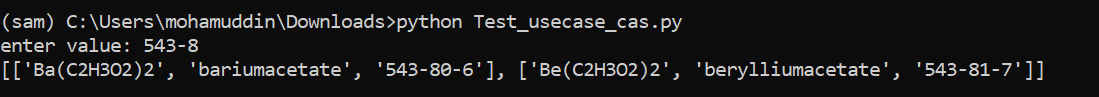
**Description:** Searches for chemical compounds using the provided CAS Registry Number.

**input\_str**: The CAS Registry Number.

**Returns Type:** list

**Returns:** Detailed information about the compound, such as chemical name, molecular formula

**Possible outputs:**



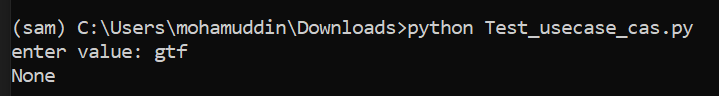
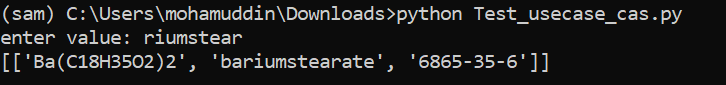
### **8. Function: search\_by\_synonyms(input\_str)**

**Description:** Searches for chemical compounds based on the provided synonym(s).

**input\_str**: The synonym or list of synonyms.

**Returns:** Detailed information about the compound(s) corresponding to the provided synonym(s)

**Possible outputs:**

.

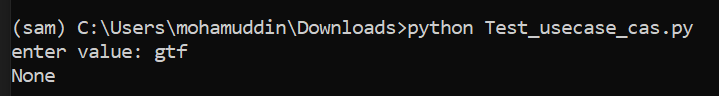
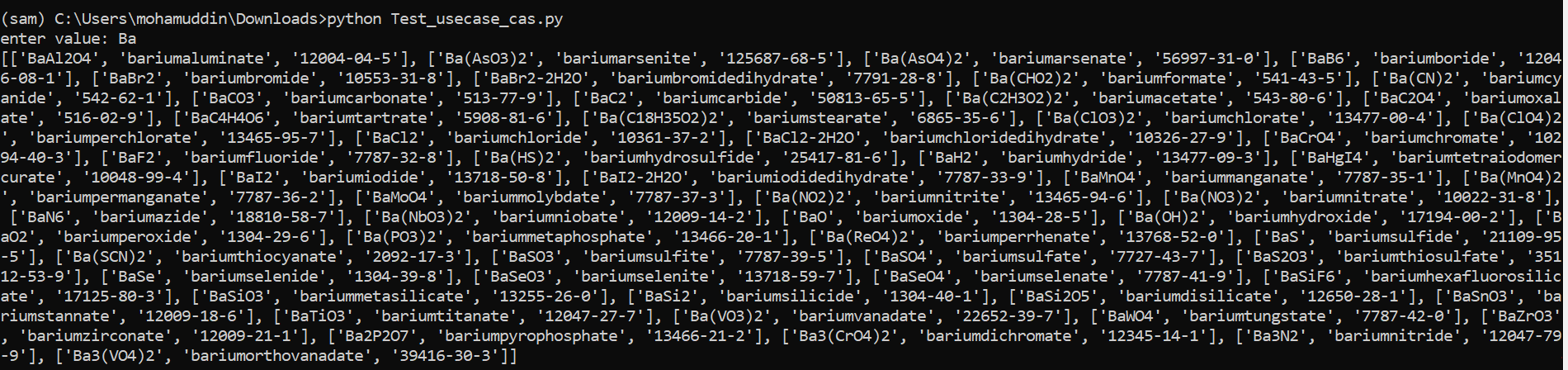
### **9. Function: search\_by\_chemicalformula(input\_str)**

**Description:** Searches for chemical compounds using the provided chemical formula.

**input\_str**: The chemical formula.

**Returns Type:** list

**Returns:** Detailed information about the compound(s) corresponding to the provided chemical formula

**Possible outputs:**.

### **10. Function: searchChemicalFormulaWithCAS(input\_str)**

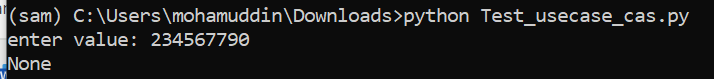
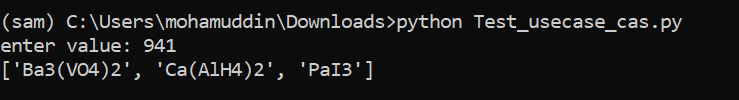
**Description:** Searches for the chemical formula associated with the provided CAS Registry Number.

**input\_str**: The CAS Registry Number.

**Returns Type:** list

**Returns:** The chemical formula corresponding to the given CAS Registry Number

**Possible outputs:**

.

### **11. Function: searchSynonymsWithCAS(input\_str)**

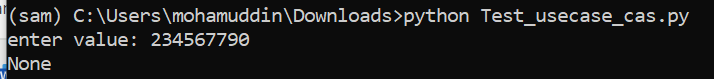
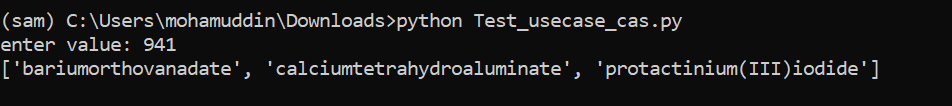
**Description:** Searches for synonyms of the compound identified by the provided CAS Registry Number.

**input\_str**: The CAS Registry Number.

**Returns Type:** list

**Returns:** Synonyms of the compound associated with the given CAS Registry Number

**Possible outputs:**

.

### **12. Function: searchCASNumberWithSynonyms(input\_str)**

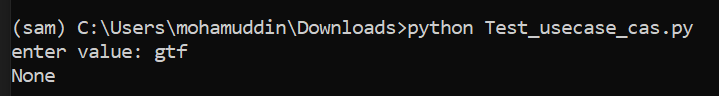
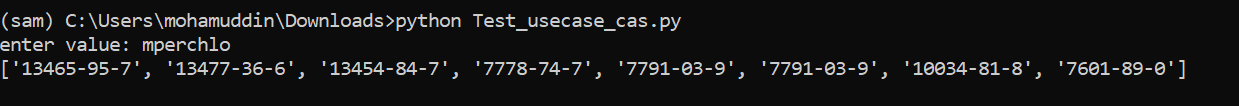
**Description:** Searches for the CAS Registry Number associated with the given synonym(s).

**input\_str:** The synonym or list of synonyms.

**Returns Type:** list

**Returns:** The CAS Registry Number(s) corresponding to the provided synonym(s)

**Possible outputs:**

.

### **13. Function: searchChemicalFormulaWithSynonyms(input\_str)**

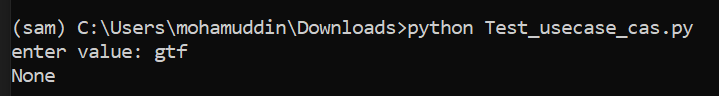
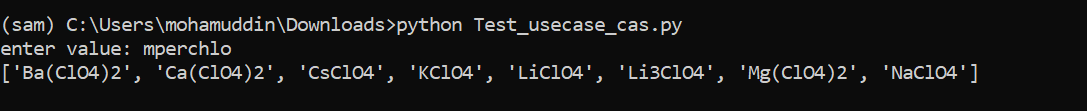
**Description:** Searches for the chemical formula based on the provided synonym(s).

**input\_str:** The synonym or list of synonyms.

**Returns Type:** list

**Returns:** The chemical formula corresponding to the provided synonym(s)

**Possible outputs:**

.

### **14. Function: searchCASNumberWithChemicalFormula(input\_str)**

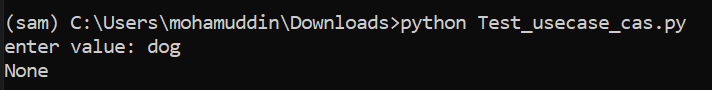
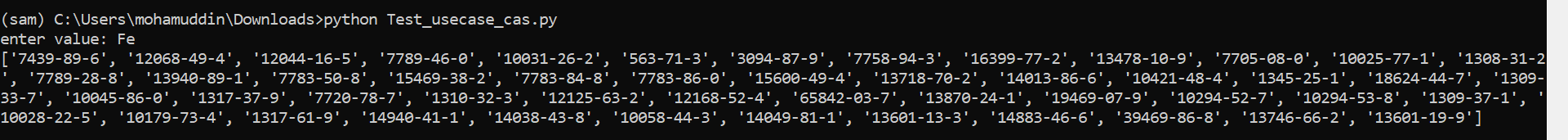
**Description:** Searches for the CAS Registry Number associated with the provided chemical formula.

**input\_str**: The chemical formula.

**Returns Type:** list

**Returns:** The CAS Registry Number corresponding to the provided chemical formula

**Possible outputs:**

.

### **15. Function: searchSynonymsWithChemicalFormula(input\_str)**

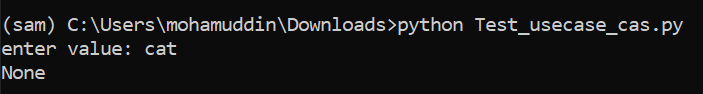
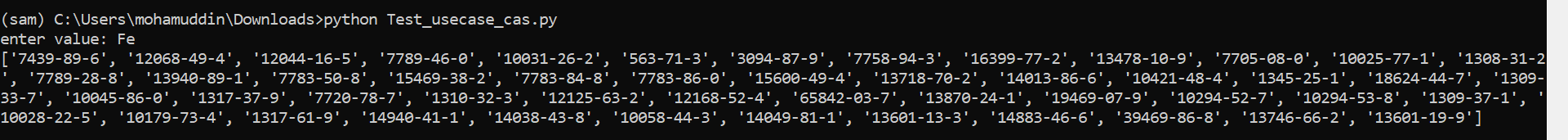
**Description:** Searches for synonyms based on the provided chemical formula.

**input\_str**: The chemical formula.

**Returns Type:** list

**Returns:** Synonyms associated with the provided chemical formula

**Possible outputs:**

.

# Python Package Creation

To publish your Python package, you typically follow these steps:

1. **Prepare Your Package:** Ensure your package is structured correctly with a setup.py file and follows Python packaging standards. Your package should include the necessary files and folders such as \_\_init\_\_.py, README.md, and any other required files.

## Package Structure: -

CAS\_RN\_Validator/

|- README.md

|- setup.py

|-src

|- cas\_rn\_validator/

|- \_\_init\_\_.py

|- validator.py ->

|-validateCAS\_RN()

|-validateCASwithMessage()

|- search.py

|- getChemicalFormulaWithCAS()

|- getSynonymsWithCAS()

|- getCASNumberWithSynonyms()

|- getChemicalFormulaWithSynonyms()

|- getCASNumberWithChemicalFormula()

|- getSynonymsWithChemicalFormula()

|- search\_by\_cas\_number(input\_str)

|- search\_by\_synonyms(input\_str)

|- search\_by\_chemicalformula(input\_str)

|- searchChemicalFormulaWithCAS(input\_str)

|- searchSynonymsWithCAS(input\_str)

|- searchCASNumberWithSynonyms(input\_str)

|- searchChemicalFormulaWithSynonyms(input\_str)

|- searchCASNumberWithChemicalFormula(input\_str)

|- searchSynonymsWithChemicalFormula(input\_str)

|- other\_module.py

1. **Set up setup.py:** Your setup.py file should contain metadata about your package such as its name, version, description, dependencies, and other relevant information.
2. **Build Your Package:** Use setuptools to create a distribution package. You can do this by running python setup.py sdist which will create a source distribution package.

* pip install setuptools
* pip install wheel
* python setup.py sdist bdist\_wheel
* pip install .\casrnvalidator-0.1.0-py3-none-any.whl

## Update Dataset

# Navigate to data folder in package CAS\_RN\_Validator\src\cas\_rn\_validator\data.

1. Update or replace the dataset.
2. Make sure data is in csv format.
3. Note: if you are replacing the dataset follow the naming convention of dataset as **cas\_rn\_dataset.csv.**
4. Navigate to parent folder in package CAS\_RN\_Validator and open command prompt.
5. Execute the command python setup.py sdist bdist\_wheel
6. Navigate to dist folder in package CAS\_RN\_Validator\dist using **cd dist** command in command prompt.
7. Execute the command pip install .\casrnvalidator-0.1.0-py3-none-any.whl

## Python code



Team Details:

Use case Proposed by: [msasidharanlalithama@deloitte.com](mailto:msasidharanlalithama@deloitte.com)

Development team: [shreybansal@deloitte.com](mailto:shreybansal@deloitte.com), [suhkhan@deloitte.com](mailto:suhkhan@deloitte.com), [mohamuddin@deloitte.com](mailto:mohamuddin@deloitte.com)