NewgenONE

Data Science Studio

Release Notes

Version: 1.1

Release Date: 11th February 2025

Disclaimer

This document contains information proprietary to Newgen Software Technologies Limited. The user may not disclose or use any proprietary information or use any part of this document without written permission from Newgen Software Technologies Limited.

Newgen Software Technologies Limited makes no representations or warranties regarding any software or the contents or use of this manual. It also specifically disclaims any express or implied warranties of merchantability, title, or fitness for any particular purpose. Even though Newgen Software Technologies Limited has tested the hardware and software and reviewed the documentation, it does not guarantee or imply that this document is error-free or accurate regarding any particular specification. As a result, this product is sold as it is, and the user, the purchaser, is assuming the entire risk as to its quality and performance. Further, Newgen Software Technologies Limited reserves the right to revise this publication and make changes in its content without any obligation to notify any person of such revisions or changes. Newgen Software Technologies Limited authorizes no Newgen agent, dealer, or employee to make any modification, extension, or addition to the above statements.

Newgen Software Technologies Limited has attempted to supply trademark information about company names, products, and services mentioned in this document. The trademarks indicated below were derived from various sources.

Copyright © 2025 **Newgen Software Technologies Ltd.** All rights reserved.
No part of this publication may be reproduced and distributed without the prior permission of Newgen Software Technologies Limited.

Newgen Software, Registered Office, New Delhi
E-44/13
Okhla Phase- II
New Delhi - 110020
INDIA
Phone: +91 1146 533 200
info@newgensoft.com

Table of contents

1 About this document 4
1.1 Revision history 4
1.2 Intended audience 4
1.3 Documentation feedback 4
2 Release summary 5
2.1 New features and enhancements 5
2.2 Removed components 6
3 Supported platforms 6
4 System requirements 6
4.1 Hardware requirements 6
4.2 Software requirements 7
5 Deployment steps 7
6 Software deliverables 7
7 Containerization support 7
7.1 Supported environments for containerization 8
7.2 Container deliverables 8
8 Constraints and limitations 9
9 Source and sink databases 9
10 Tech stack 10
11 Release artifacts 10
12 Version number 10
13 Release dates 10
14 Product development location 11

1 About this document

These release notes provide information on the features, supported platforms, and system requirements of NewgenONE Data Science Studio 1.1.

1.1 Revision history

Revision date	Description
February 2025	Initial publication

1.2 Intended audience

These release notes are intended for the following:

- Data engineers
- Business analysts
- Citizen and expert data scientists

1.3 Documentation feedback

To provide feedback or any improvement suggestions on technical documentation, write an email to docs.feedback@newgensoft.com.

To help capture your feedback effectively, share the following information in your email:

- Document name
- Version
- Chapter, topic, or section
- Feedback or suggestions

2 Release summary

This release of the NewgenONE Data Science Studio platform supports on-premises deployment in both Docker and non-Docker environments. Application sign-in is facilitated through Microsoft's Active Directory using Lightweight Directory Access Protocol (LDAP) for on-premises deployment.

2.1 New features and enhancements

The following table lists the new features category and components:

S. No.	Component	Additions
1	Algorithms	• Dex on Time Series data • MLP Regression Algorithm • Light GBM • XG-Boost • SVM for regression • KNN clustering • HDB Scan clustering • DB Scan clustering
2	Data source	• NewgenONE source for reading data from OmniDocs • PostgreSQL • Oracle
3	Predictive Model Markup Language (PMML) API	PMML API details are available directly in the production environment.
4	GenAI	Data exploration, Dex actions, and SHAP (SHapley Additive exPlanations) values are now accessible through simple queries in GenAI.

The following table lists the improvements in NewgenONE Data Science Studio 1.1:

S. No.	Component	Improvements
1	Pipeline monitoring	The pipeline monitoring screen includes an option to view pipeline results.
2	Feature tool	FeatureTool datasource now includes Oracle, Microsoft SQL, and PostgreSQL.
3	Dev – Serving separation	Separate environments for development and serving pipeline retraining. The PMML Inference API is available on serving.
4	Model metrics	Model monitoring metrics include CSI values, KS Test, Lift, and GAI.
5	SMOTE	• Synthetic Minority Over-sampling Technique (SMOTE) supports data sampling for categorical columns. • An under-sampling option is provided for data sampling.

2.2 Removed components

The following components and features are removed in NewgenONE Data Science Studio 1.1

- Auto Dex
- Drools rule engine
- Elastic search sink
- SHAP clickable result

3 Supported platforms

NewgenONE Data Science Studio 1.1 supports the following environments:

Browser	Language supported
• Microsoft Edge 131.0 and later • Chrome 131.0 and later • Mozilla Firefox 133 and later	English (United States)

4 System requirements

This section describes the minimum hardware and software requirements for using the NewgenONE Data Science Studio 1.1 platform.

4.1 Hardware requirements

The table below lists the minimum on-premises resource requirements:

Resource name	Quantity	Details
CPU	3	• Cores—32 virtual cores • RAM—128 GB • Operating system disk—128 GB • Hard drive—1024 GB
GPU	1	• CPU ➢ Cores—8 (16 v-cores) ➢ RAM—32 GB • Processor—Intel(R) i7-11700K • GPU—RTX 5000 ➢ RAM—24 GB RAM ➢ CUDA cores—8,192 ➢ NVIDIA Tensor cores—256 • CUDA ➢ Version—12.4 or later ➢ Driver version—550.54.14 or later

Resource name	Quantity	Details
		• Storage ➢ SSD OS drive—256 GB space ➢ SSD external drive—512 GB space

4.2 Software requirements

This section lists the minimum software requirements for using the NewgenONE Data Science Studio 1.1 platform.

- **Operating system on virtual machine**—RHEL 8.9
- **Third-party software** (required for Docker deployments only)—Docker 27.5 or later version, Docker Compose 2.32 or later version.

5 Deployment steps

The deployment process is executed manually through scripts. For procedural details on how to perform the deployment, refer to the product installation and configuration guides.

6 Software deliverables

The following are the software deliverables:

- Docker images for docker deployment.
- The ZIP packages are used for the application in standard deployment within a non-Docker environment.

7 Containerization support

This section describes the supported environments and deliverables of Docker containerization of NewgenONE Data Science Studio 1.1.

7.1 Supported environments for containerization

The table below lists the combination of the supported environments for the installation of NewgenONE Data Science Studio on-premises:

Supported environment	On-premises
Container orchestration	• Docker Compose 2.32.51 or later • Docker 27.0.1 or later The following are the supported sub-platforms: • Host operating system—RHEL 8.9 • Container base image—RHEL 8.9 • Application server—Node 18.19.x
Database server	PostgreSQL 15.3 container image
Language	English (United States)

7.2 Container deliverables

The table below lists the deliverables for the on-prem container environment:

Deliverables	On-premises
Docker images	• NAI User Instance Web for Model Development • NAI User Instance Web for Model Production • PostgreSQL • NAI backend Modelling API service • NAI backend Serving API service • NAI backend JWT token generator API service • JupyterHub Service • Upload document API service • PMML API service • Backend Gen AI API service • Gen AI Dex API service • Gen AI Dex-action API service • Gen AI serving API service • Gen AI Ollama server service
Configuration files	• For Gena AI • For web Application • For backend application • Lemma config for backend
YAML files	Docker-compose.yml
Deployment and configuration documents	NewgenONE Data Science Studio Installation and Configuration Guide for Docker

8 Constraints and limitations

This release of NewgenONE Data Science Studio has the following constraints and limitations:

- One Hot Encoding (OHE) and String Indexer (SI) cannot be used together.
- The ID column for inference must be a non-feature column in the dev pipeline's data source.
- Users can schedule retraining pipelines in the production environment.
- Only assigned approvers can approve pipelines for production.
- Missing value imputation applies to null values, not empty values.
- Characteristic Stability Index (CSI) graphs are unavailable if Principal Component Analysis (PCA) or Singular Value Decomposition (SVD) algorithms are used in the model.
- Alternating Least Squares (ALS) recommendations for new users in the data requires a model rebuild.
- XGBClassifier requires labels to start from 0 in multiclass classification.
- Serving is not supported for the following algorithms:
  - Affinity
  - Frequent Pattern Growth (FPG)
  - Accelerated Failure Time (AFT)
  - TriangleCount
  - Page Rank
  - Latent Dirichlet Allocation (LDA)
  - Word2Cluster (W2C)
  - Arima
  - Arimax
  - AutoArima
- Oracle data source and sink connections without SSL are supported only in Docker deployment.
- Oracle source works with a single schema only.
- Model development pipelines without an evaluator cannot be published to production.

9 Source and sink databases

The table below lists the recommended database versions for use as data sources and/or sinks in pipelines:

Database	Recommended version
MySQL	8.0–8.4
Microsoft SQL	SQL Server 2014 (12.x)
PostgreSQL	15.3–17.2
Oracle Database	19c and 21c

10 Tech stack

- React—16.13.1
- Node.js—18.19.0
- Python—3.11.10
- PostgreSQL—15.3

11 Release artifacts

The release artifacts are as follows:

- Deployment guide
- Release notes

S. No.	Enclosed	Yes / No / NA
1.	Installation guide (describing installation procedure for the network as well as standalone version; dependency of software on the operating system and other software; structure and paths of various *.*ini* files, minimum software, and hardware requirements)	Yes
2.	New function listing	Yes
3.	Modified functions	NA
4.	Interface modifications	NA
5.	Problem status (the bug IDs and descriptions for all the pending bugs)	NA
6.	COTS components (if any)	NA
7.	Rollback steps (in the case of a separate document)	Yes
8.	Deployment steps (in the case of a separate document)	Yes
9.	Others (if any, for example - freeware, open-source, and executables).	Yes

12 Version number

NewgenONE Data Science Studio 1.1

13 Release dates

Following are the release dates:

- **Product release date**—11th February 2025
- **Market launch date**—14th February 2025

14 Product development location

Newgen Software Technologies Limited,
10th & 12th Floor, Tower No.11,
IT/ITES SEZ, Candor TechSpace,
Plot No.20 & 21, Sector-135,
Noida, Uttar Pradesh – 201304
GSTIN No. - 09AAACN0500F4Z1
E-mail - info@newgensoft.com