

Neuton Quick Start Guide



____ CONTENT

Log in to Neuton	3
Select data for training	4
Train your model	8
Make predictions1	0

LOG IN TO NEUTON

Go to https://neuton.ai in your web browser to access Neuton. After page refresh click **«Sign in»** to start using Neuton. You will be redirected to the **«Sign in with Google»** web page, after that please provide the **email** address and **password** used during Neuton registration.



SELECT DATA FOR TRAINING

Create your first solution using the **«Add Solution»** button.



Specify the **«Solution Name»** in the pop-up window.

Intelligent Agont Neuton	reese Google Cloud	THIRD LAW 🐥 Isaac Newton 👻 🕐
🕅 My Solutions		
Business Use Cases		
😂 Dataset Storage		
G Google Cloud	New Solution 🗶	
D Support Library		
💭 Quick Start Tour	Solution Name Example_Ttanic 25 Character Limit 15/25 Description 140 Character Limit 0/140 Cancel Next →	
\$ Eligible for an additional \$200 in credits? Get it now! Learn more		

Press the **«Next»** button to select the preloaded dataset or upload a new one.



On the next screen press the **«Select Dataset»** button. Upload training dataset from the Kaggle Titanic ML competition or navigate to the **«Preloaded dataset»** tab and choose **Titanic**:

Neuton	Course Google Cloud Example_Titanic ver. 1	THIRD LAW	🔎 Isaac Newton 👻 ?
My Solutions	Dataset Training Prediction		Advanced mode
Business Use Cases			
😂 Dataset Storage			
G Google Cloud	Training Dataset	×	
Support Library			
D Quick Start Tour	Vell known Datasets Vell known Datasets O Titanic (Binary Classification) O House Prices (Regression) O Ghouls, Goblins, and Ghosts (Multiclassification) Retail D Finance OK Dataset Requirements		
\$ Eligible for an additional \$200 in credits? Get it now! Learn more		l	Eel Integrator & 2127

Press the **«OK»** button to confirm selected dataset.

NOTE:

List of requirements for training datasets in Neuton

- 1. Dataset must be a CSV file using UTF-8 or ISO-8859-1 encoding.
- 2. Dataset must have a minimum of 2 columns or rows.
- 3. Dataset must not have any empty columns.
- 4. Pay attention to duplicate rows: Duplicate rows will be removed from the training set during the preprocessing stage, which may cause an inconsistency with the dataset minimum size and sample number requirements.
- 5. The first row in the dataset must contain the column names, and a comma, semicolon, pipe, caret or tab must be used as a separator. CRLF or LF should be used as the end of a line character. The separator and end of line character should be consistent inside the dataset.
- 6. All column names must be unique and must contain only letters (a-z, A-Z), numbers (0-9), hyphens (-) or underscores (_).
- 7. All empty values must be excluded from the target variable.
- 8. For regression task types the target variable must have only numeric values.
- 9. For classification task types at least ten samples must be provided for each class of target variable.



- 10. Currently Neuton supports only the EN-US locale for numbers and timestamps, so:
 - You must use a dot as a decimal separator, and delete spaces and commas typically used to separate every third digit in your numeric fields.
 - Example "20,000.00" should be replaced with "20000.00"
 - If any numeric column is represented as a combination of a number and its corresponding unit, then only the number should be placed in the column.
 - Example "\$20,000.00" should be replaced with "20000.00"
 - You can use any of the following date/timestamp formats:
 - "DD.MM.YYYY hh:mm:ss"
 - "MM.DD.YYYY hh:mm:ss"
 - "YYYY.MM.DD hh:mm:ss"
 - "YYYY.DD.MM hh:mm:ss"
 - "DD.MM hh:mm:ss"
 - "MM.DD hh:mm:ss"
 - "MM.YYYY hh:mm:ss"
 - "YYYY.MM hh:mm:ss"

You can use "/",":" or "." as date components separator. Space should be used as date and time separator and ":" should be used as time components separator.

- The timestamp format should be consistent inside any given column.
- In test datasets the same timestamp format should be used in a manner consistent with the training dataset.
- 11. Text fields should start and end with quotes, in order to avoid confusing symbols in the text with column separators.
- 12. End of line symbols must be excluded from the fields values.
- Neuton supports NLP, but can currently only process textual fields in English (US). If you used a language different from English for categorical fields in your dataset, you must transliterate them into English (US).

List of requirements for test datasets (or new data)

- 1. Dataset must be a CSV file using UTF-8 or ISO-8859-1 encoding.
- 2. The first row in the dataset must contain the column names, and a comma, semicolon, pipe, caret or tab must be used as a separator. CRLF or LF should be used as the end of a line character.
- 3. The test dataset must have the same file structure and data types as the training dataset.



- 4. The column names must be the same and the columns must have the same order as in the training dataset.
- 5. The target variable column and its name must be excluded from the test dataset.
- 6. For the Gravity Subscription Plan the size of your test set should not exceed 100 rows.
- 7. In test datasets the same timestamp format should be used in a manner consistent with the training dataset.
- 8. Text fields should start and end with quotes, in order to avoid confusing symbols in the text with column separators.
- 9. End of line symbols must be excluded from the fields values.

Select Target Variable.

Intelligent Agent Newton	Petter Google Cloud Example_Titanic ver.	1	THIRD LAW 🔺 Isaac Newton 👻 ?
My Solutions	Dataset	Training Prediction	Advanced mode
(ff) Business Use Cases			
😂 Dataset Storage		Training Dataset	
G Google Cloud		Q Titanic (Binary Classification) Change	
Support Library		- proceeded outdates realing-anning-an	
Duick Start Tour		Target Variable	
\$ Eligible for an additional \$200 in credits? Get it now!		Select target variable Q ● Survived ● ● Polass ● ● Name ● ● Sex ● ● Age ●	
Learn more			Bell Integrator © 2021

Press the **«Next»** button to go to next step.

NOTE:

For the preloaded datasets all settings are preconfigured.



TRAIN YOUR MODEL

On this step **Neuton** will automatically offer you **Task type** and target **Metric** but you can change it according to your preferences. To continue press **«Start Training»**.

Intelligent Agent Newton	Paster Google Cloud Example_Titanic ve	r.1 T	HIRD LAW A Isaac Newton 👻 ?
😚 My Solutions	Dataset	Training Prediction	Advanced mode
Business Use Cases			
😂 Dataset Storage		Training dataset: Titanic (Binary Classification	
G Google Cloud		Target variable: Survived	
Support Library		Task type:	
도 Quick Start Tour		Binary Classification V Metric: Accuracy V TinyML	
\$ Eligible for an additional \$200 in credits?			

You can track the percentage of training completion on the progress bar at the top of the training tab.

Intelligent Agent Newton	Paster Google Cloud Exa	nple_Titanic ver. 1	THIRD LAW	🌲 Isaac Newton 👻 ?
😚 My Solutions	Dataset	Training	Prediction	Advanced mode
Business Use Cases				
😂 Dataset Storage		Status Complete Virtual machine creation		
G Google Cloud		Not consistent	محــ	
D Support Library		Progress		
Quick Start Tour				
		Target metric: Accuracy	Analytic's Tools	
		Task type: Binary Classification Target variable: Survived	Training dataset: Titanic (Binary Classific	
		Metrics	? ~	
\$		Log	~	
Eligible for an additional \$200 in credits? Get it now!				
Learn more				Bell Integrator © 2021

As a result of training, the model is created.





Press the **«OK»** button to go to Prediction stage.



MAKE PREDICTIONS

When the training is completed click the **«Enable»** button to switch on prediction.

Neuton	Petter Google Cloud Example_Titanic ver. 1			THIRD LAW Isaac Newton 👻 ?
😧 My Solutions	Dataset	Training	Prediction	Advanced mode
Business Use Cases				
😂 Dataset Storage	Historical Model-to-Data Relevance	Indicator	Prediction	ර් Enable
G Google Cloud		4	<u>^</u>	
D Support Library	To use this tool, first enable in the Prediction Settings ar	Model-to-Data Relevance Indicator d make a prediction on new data.	Web Prediction	Start
Quick Start Tour	~	- T		
			REST API Access	Access Instructions
	Target metric: Accuracy 0.836158	Analytic's Too	bls	
	Task type: Binary Classification	Training dataset: Titanic (Binary Classi	fic	n 🛓 Download
	Target variable: Survived	Q Processed dataset: processed_trai	n.csv	
	Metrics	?	~	
<u> </u>	Prediction Settings	i.	^	
Eligible for an	Model to Data Relevance Indicator			
additional \$200 in credits?				
Get it now!				•
Learn more				Bell Integrator © 2021

NOTE:

That additional costs may be charged for a new virtual machine for predictions.

When the Virtual machine is created, the **«Start»** button will be activated. Press this button to start **Web Prediction**. In the pop-up window click **«Start»** (for the preloaded datasets the test dataset is preselected). If you haven `t used the preloaded dataset you can upload your test dataset (or new data) without a target variable to make the prediction (this process is similar to training dataset selection). You can download a test dataset from the Kaggle Titanic ML competition.

NOTE:

The test dataset structure and variable types should be the same as in the training dataset, except that the target column will be excluded.



Intelligent Agent Neuton	Partor Google Cloud Example_Titanic ver.	I	THIR	o Law isaac Newton 👻 ?
My Solutions	Dataset	Training	Prediction	
Business Use Cases				
😂 Dataset Storage	Historical Model-to-Data Relevand	e Indicator	Prediction	() Disable
G Google Cloud		2 de la compañía	Mah Drediction	Cana
Support Library	To use this tool, first enal in the Prediction Setting	Ne Model-to-Data Relevance Indicator	web Prediction	Start
Quick Start Tour	Predict	ing on your data	Target variable: Survived	Access Instructions
	Target metric: Accuracy Progress		Duration 00 07	-
	Task type: Binary Classification	Stop		🛓 Download
	Target variable: Survived		-	
		processed_train.c	sv	
é	Metrics	? ~		
P Eligible for an	Prediction Settings	^		
additional \$200 in credits?	Model-to-Data Relevance Indicator			
Get it now!				6
Learn more				Bell Integrator © 2021

Wait for predicting completion. It can take a few minutes.

Congratulations! Your predictions are completed!

Neuton	Parter Google Cloud Example_Titanic ver. 1			THIRD LAW Saac Newton 👻 ?
😚 My Solutions	Dataset	Training	Prediction	
聞 Business Use Cases				
😂 Dataset Storage	Historical Model-to-Data Relevance Ind	icator	Prediction	U Disable
G Google Cloud		1 × 1		Case
D Support Library	100%	Prediction completed	× nic_v.1	Start
Puick Start Tour		5.	02 PM	٩
	Target metric: Accuracy 0.836158		ess	Access Instructions
	Task type: Binary Classification	T Prediction results are ready for Exam	nple_Titanic	
	Target variable: Survived	🛓 Download 🔍 V	le solution	🛓 Download
	Metrics	? ~		
Fligible for an	Prediction Settings	^		
additional \$200 in credits?	Model-to-Data Relevance Indicator			
Get it now!				6
Learn more				Bell Integrator © 2021

Download prediction results for further insights!

For more information about Data Science and improving prediction results on your dataset please refer to <u>Neuton User Guide</u>.

