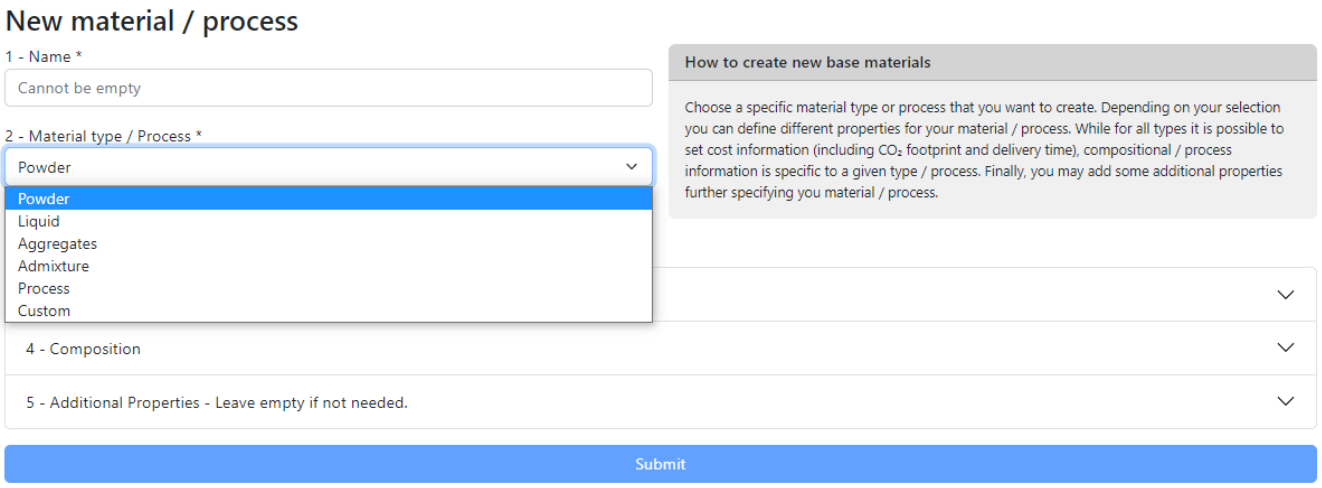
User Manual

Based on the functionality provided with SLAMD, this document provides a guide of how to use the app.

The general workflow is as follows. First define base materials and processes. Then, you use the former as a basis to create blended materials. Next, a materials
formulations can be specified. For this purpose one selects a subset of all materials and processes created. The resulting data can but must not be enriched with targets (labels).
This specification of data can now used in the sequential learning for the prediction of new material properties.

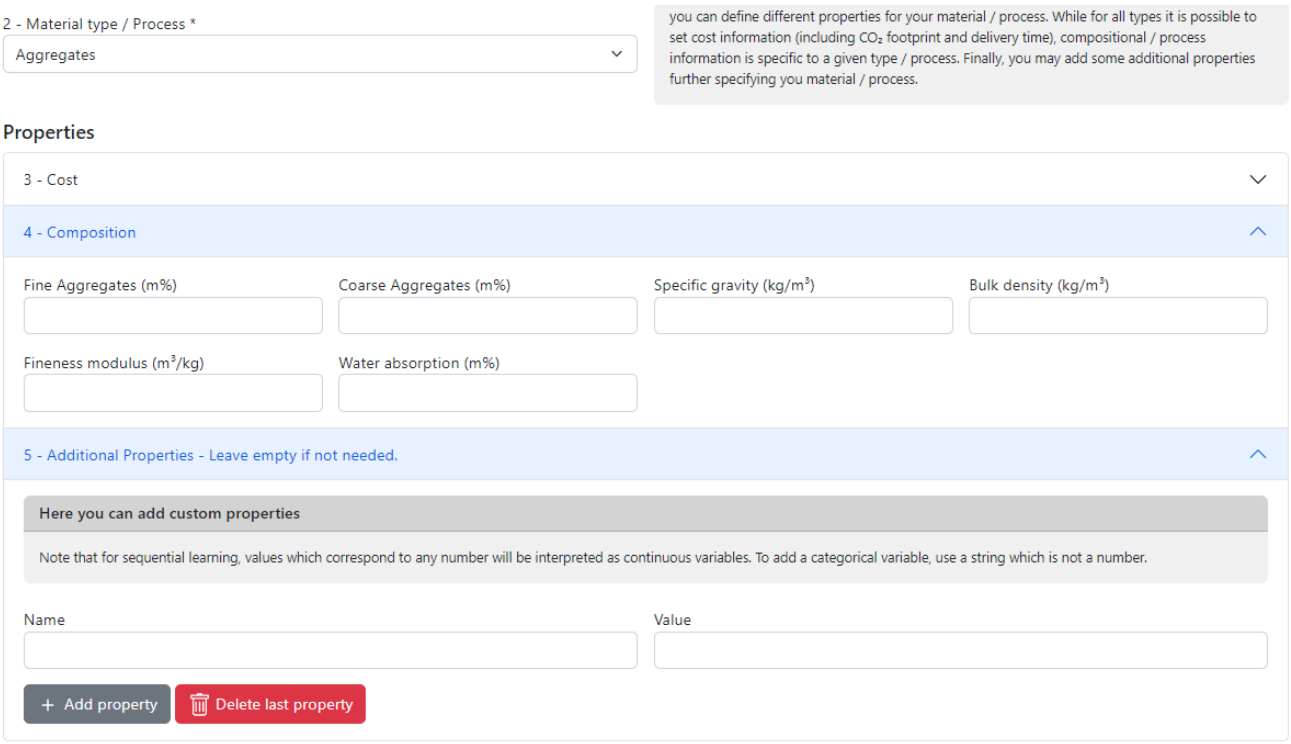
1. Base Materials and Processes

The starting point is the creation of base materials and processes. Here you can first select the type of base material or process you want to specify.



Note that for all the types you can specify a name, costs and additional properties in the same way. The concrete composition, however, depends on the type (e.g. in the screenshow below, powders is shown.
Furthermore, there is the possibility to specify custom materials. These are not restricted and can in principle describe any type of material. As such, it has no a priori composition and must thus be specified in terms
of its additional properties solely. Note that the latter can be used to specify a custom composition anyway.

You can open the cost / composition / additional properties input fields by clicking on the corresponding item below "Properties".



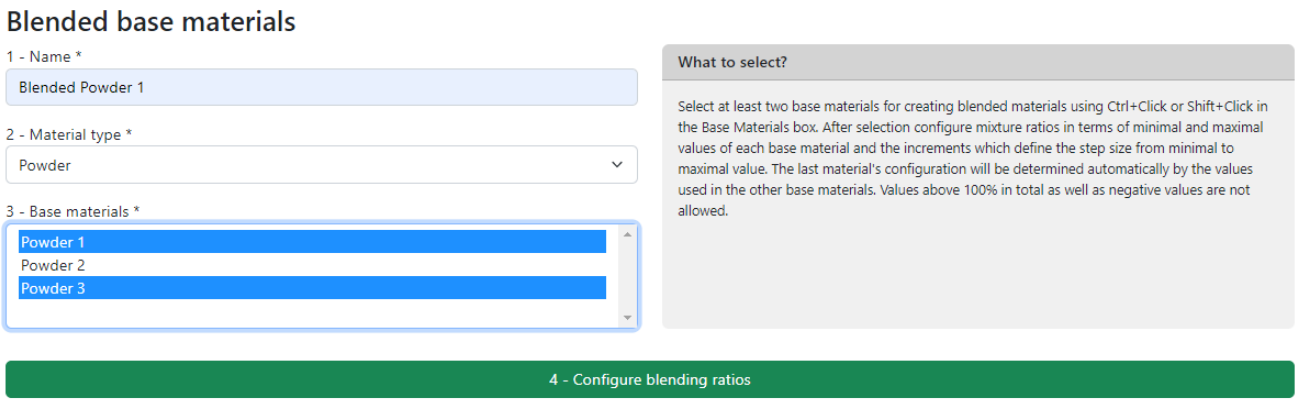
Depending on the field, either only numerical or any alphanumeric input can be specified. Note that for later creation of blended materials, a base materials with
at least one field empty (in costs or composition) is considered incomplete and a warning is shown. Nevertheless blending can be still be performed as one might not want to create base materials
and blends consisting of all the possible features. More details on blending rules will be specified below.

The additional properties (at most 10!) allow defining additonal custom features for the material. In principle, you are free to specify any name and corrsponding value. Note however, that when blending
there are again certain rules which apply. For a complete description the names and corresponding value types must match across all base materials used for blending.

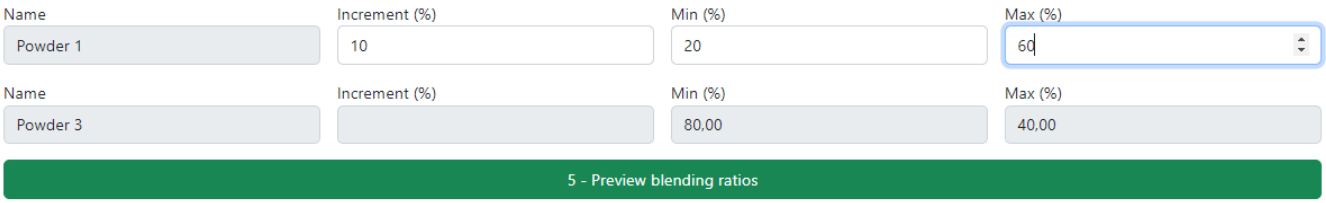
Having configured your material you can save it by clicking on the "Save material" button. The new item appears in the table showing all base materials and processes. In the left column you have functionalities
to delete or edit the chosen material. In case of clicking the edit button, the form will be populated by the previously configured data.

2. Blended Materials

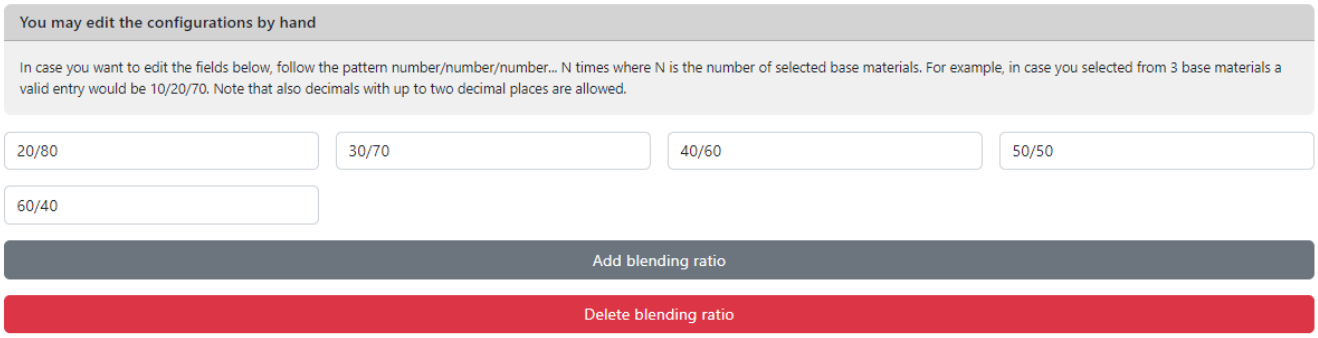
The next step is creating blended materials. As a starting point set a name (mandatory) and a type and choose at least two base materials:



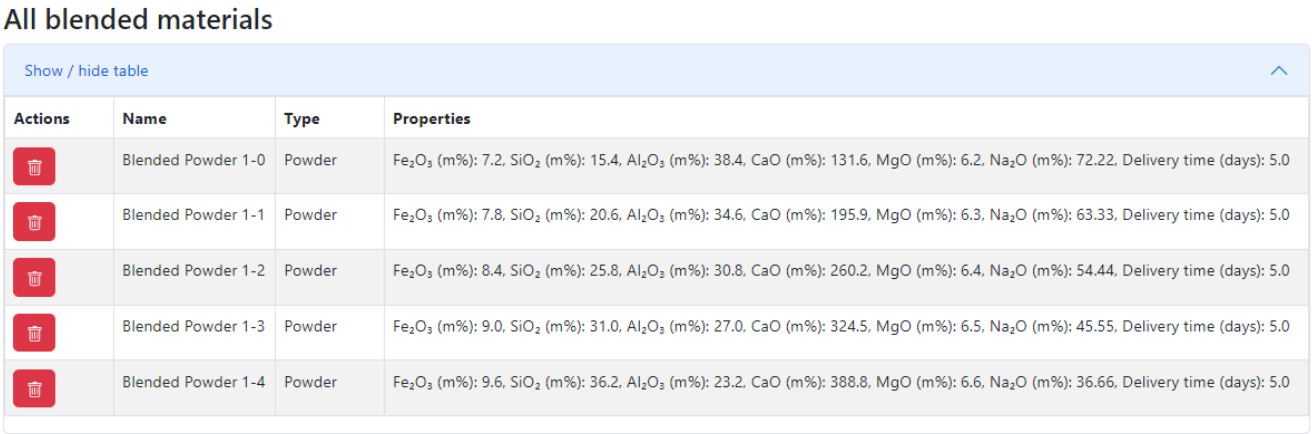
Blending then basically corresponds to mixing different materials of a given type with one another to create materials with various ratios of the properties for all base materials involved.
For example, in case of the hypothetical situation where you have two powders both solely described by 'Specific gravity' you can define a blend e.g. of 10% the first powder and 90% the second one. The resulting Specific gravity
is then computed as the weighted sum of the original Specific gravities. Note that the same logic applies to all the other properties except for 'Delivery time' which is defined as the maximum of the base materials Delivery times.



The relative ratios can be specified after having selected the base materials used for blending. Note that the last row is always determined by the previous ones. In the screenshot, e.g. we blend from two base admixtures (in general, we can use more than just two).
The min and max values of the last material are determined by the constrained that the sum over all has to equal 100%. Accordingly, one can only specify values and increments for the independent materials. The resulting ratios are computed as the direct product of all possibilities, e.g. in the scenario shown in the above image we get



If we were to submit this configuration, we would generate five blended admixtures. Note, however, that it is still possible to add, delete and edit the mixtures. Here, there is no more restriction that everything needs to add up to 100. However, one must follow the correct pattern when editing (as shown in the comment in the screenshot).
Finally, once you are finished with the specification of the blending configuration, click the "Create blended materials" button to create your blends. The created entries are saved and displayed in a table which show all details (here we show some blended materials which we have already created):



In addition, there are several rules which apply when not all fields are filled in the base materials or in case some properties are only filled for one of the base materials to be mixed.

- A description is complete only when all input fields are filled (except for additional properties)

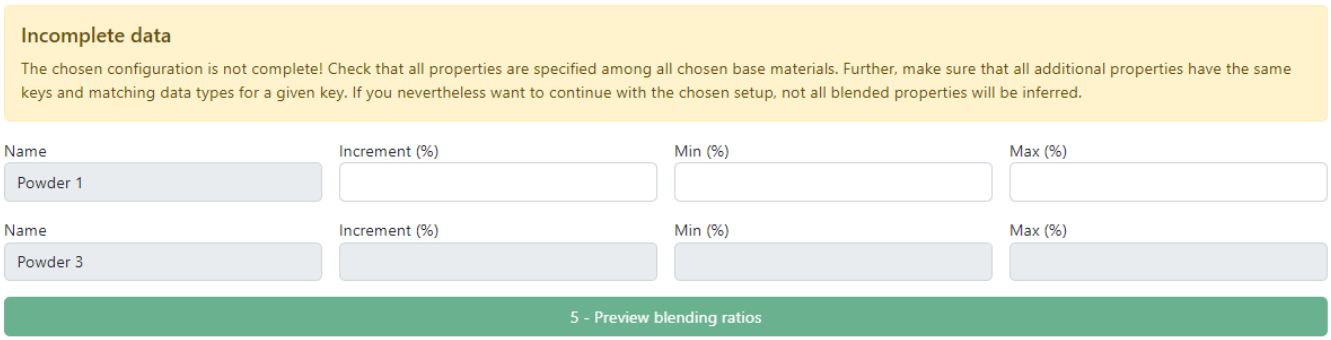
- Assume you want to mix two (or more) base materials which both define a compositional / costs or structural property which we call 'Base Material Property' in the following (it could e.g. be Specific Gravity in case of powders)
  and assume both define an additional property calles 'Add. Prop'. Then the following rules apply when blending:

  a) Base Material Property is filled for all base materials the resulting blended materials have the property Base Material Property computed as the weighted sum of its base materials Base Material Property values
  b) At least one of the base materials did not specify a value for Base Material Property the resulting blend will not have the property either as the description is not complete
  c) Add. Prop is filled for all base materials and has only numeric value the resulting blended materials have the property Add. Prop computed as the weighted sum of its base materials Add. Prop values
  d) Add. Prop is filled for all base materials and has only non-numeric values a new Property for all differing values of the corresponding base materials Add. Prop will be created with the value set by the relative weight of the base material. E.g if we have two powders,
  Powder A and Powder B and Add. Prop of Powder A is set as X and Add. Prop of Powder B is set as Y and we have a mixing ratio of 20/80 then the resulting blend does not have an additional property Add. Prop but rather two additional properties X and Y with values 0.2
  and 0.8, respectively. This is also illustrated in the images below for the additional property 'Prop 2' (here, we have two base materials and the weights configured such that three blended materials are created from these).
  e) Add. Prop is filled for all base materials and some have non-numeric values while for others the values are numeric Incomplete description: no corresponding additional property is created for the blended materials

	Powder 1	Powder	Fe_2O_3 (m%): 12.0, SiO_2 (m%): 57.0, Al_2O_3 (m%): 8.0, CaO (m%): 646.0, MgO (m%): 7.0, Na_2O (m%): 1.1, K_2O (m%): 57.0, SO_3 (m%): 8.0, TiO_2 (m%): 5.0, P_2O_5 (m%): 6.0, SrO (m%): 4.0, Mn_2O_3 (m%): 9.1, LOI (m%): 57.0, Fine modules (m^2/kg): 68.0, Specific gravity (m%): 8.0, Costs (€/kg for materials, € for processes): 56.1, CO_2 footprint (kg/ton for materials, kg for processes): 12.0, Delivery time (days): 5.0, Prop 1: 12
	Powder 2	Powder	Fe_2O_3 (m%): 6.0, SiO_2 (m%): 5.0, Al_2O_3 (m%): 43.0, CaO (m%): 8.0, MgO (m%): 5.0, Na_2O (m%): 3.0, K_2O (m%): 10.0, SO_3 (m%): 6.1, TiO_2 (m%): 4.0, P_2O_5 (m%): 78.0, SrO (m%): 7.0, Mn_2O_3 (m%): 64.0, LOI (m%): 4.0, Fine modules (m^2/kg): 7.0, Specific gravity (m%): 4.0, Costs (€/kg for materials, € for processes): 85.0, CO_2 footprint (kg/ton for materials, kg for processes): 56.0, Delivery time (days): 4.0, Prop 1: 18

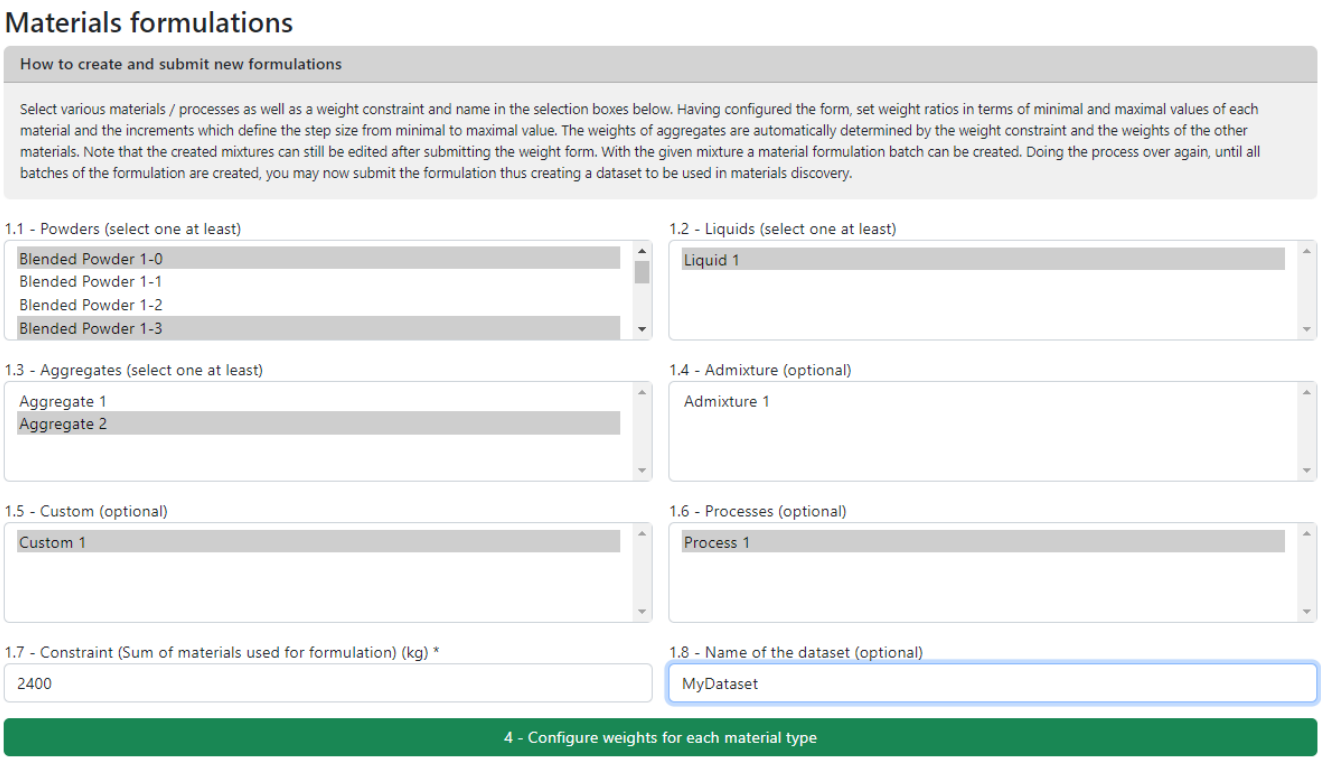
	Complete Blended Powder-0	Powder	Fe_2O_3 (m%): 8.4, SiO_2 (m%): 25.8, Al_2O_3 (m%): 29.0, CaO (m%): 263.2, MgO (m%): 5.8, Na_2O (m%): 2.24, K_2O (m%): 28.8, SO_3 (m%): 6.86, TiO_2 (m%): 4.4, P_2O_5 (m%): 49.2, SrO (m%): 5.8, Mn_2O_3 (m%): 42.04, LOI (m%): 25.2, Fine modules (m^2/kg): 31.4, Specific gravity (m%): 5.6, Costs (€/kg for materials, € for processes): 73.44, CO_2 footprint (kg/ton for materials, kg for processes): 38.4, Delivery time (days): 5.0, Prop 1: 15.6
	Complete Blended Powder-1	Powder	Fe_2O_3 (m%): 8.7, SiO_2 (m%): 28.4, Al_2O_3 (m%): 27.25, CaO (m%): 295.1, MgO (m%): 5.9, Na_2O (m%): 2.15, K_2O (m%): 31.15, SO_3 (m%): 6.96, TiO_2 (m%): 4.45, P_2O_5 (m%): 45.6, SrO (m%): 5.65, Mn_2O_3 (m%): 39.3, LOI (m%): 27.85, Fine modules (m^2/kg): 34.45, Specific gravity (m%): 5.8, Costs (€/kg for materials, € for processes): 72.0, CO_2 footprint (kg/ton for materials, kg for processes): 36.2, Delivery time (days): 5.0, Prop 1: 15.3
	Complete Blended Powder-2	Powder	Fe_2O_3 (m%): 9.0, SiO_2 (m%): 31.0, Al_2O_3 (m%): 25.5, CaO (m%): 327.0, MgO (m%): 6.0, Na_2O (m%): 2.05, K_2O (m%): 33.5, SO_3 (m%): 7.05, TiO_2 (m%): 4.5, P_2O_5 (m%): 42.0, SrO (m%): 5.5, Mn_2O_3 (m%): 36.55, LOI (m%): 30.5, Fine modules (m^2/kg): 37.5, Specific gravity (m%): 6.0, Costs (€/kg for materials, € for processes): 70.55, CO_2 footprint (kg/ton for materials, kg for processes): 34.0, Delivery time (days): 5.0, Prop 1: 15.0

Note that whenever a configuration of base materials is chosen leading any of the described incomplete scenarios, a warning is shown.



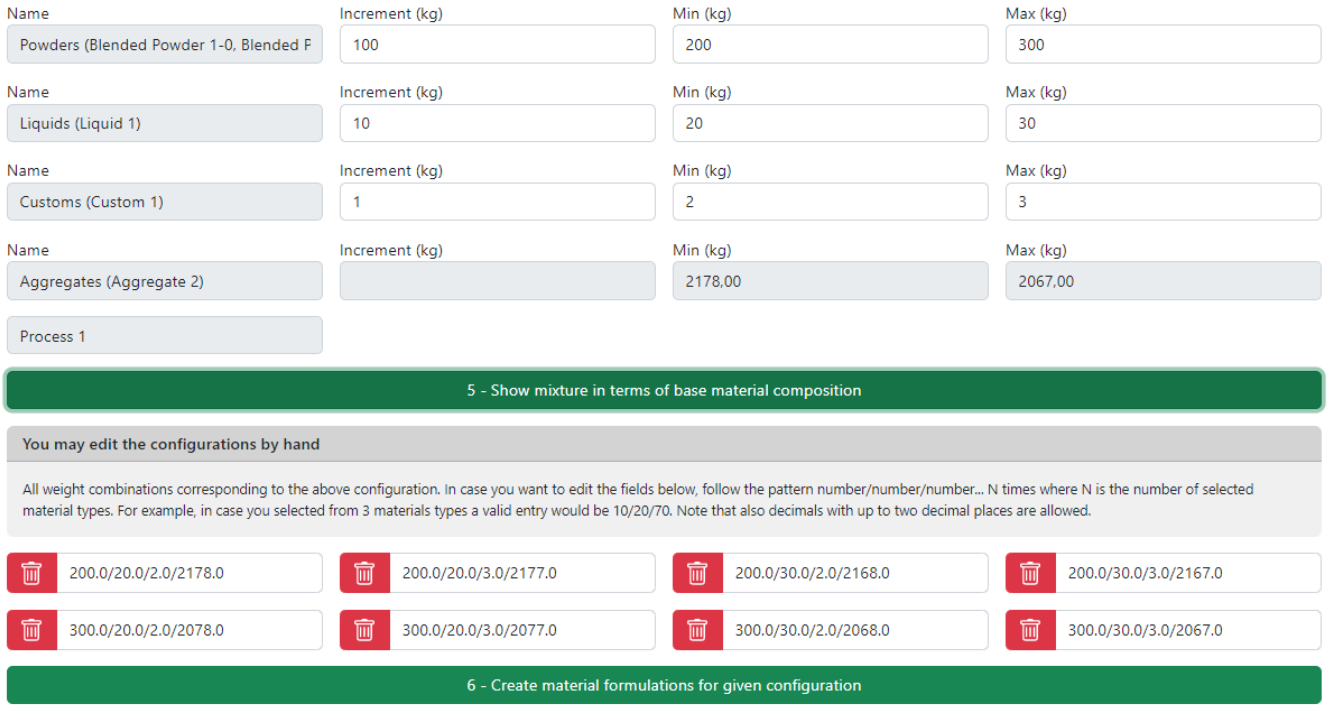
3. Material Formulations

A formulation combines various base materials and processes. Thus, one can choose multiple entries from each the materials (base as well as blended) and processes created so far. This is illustrated in the following screenshot:



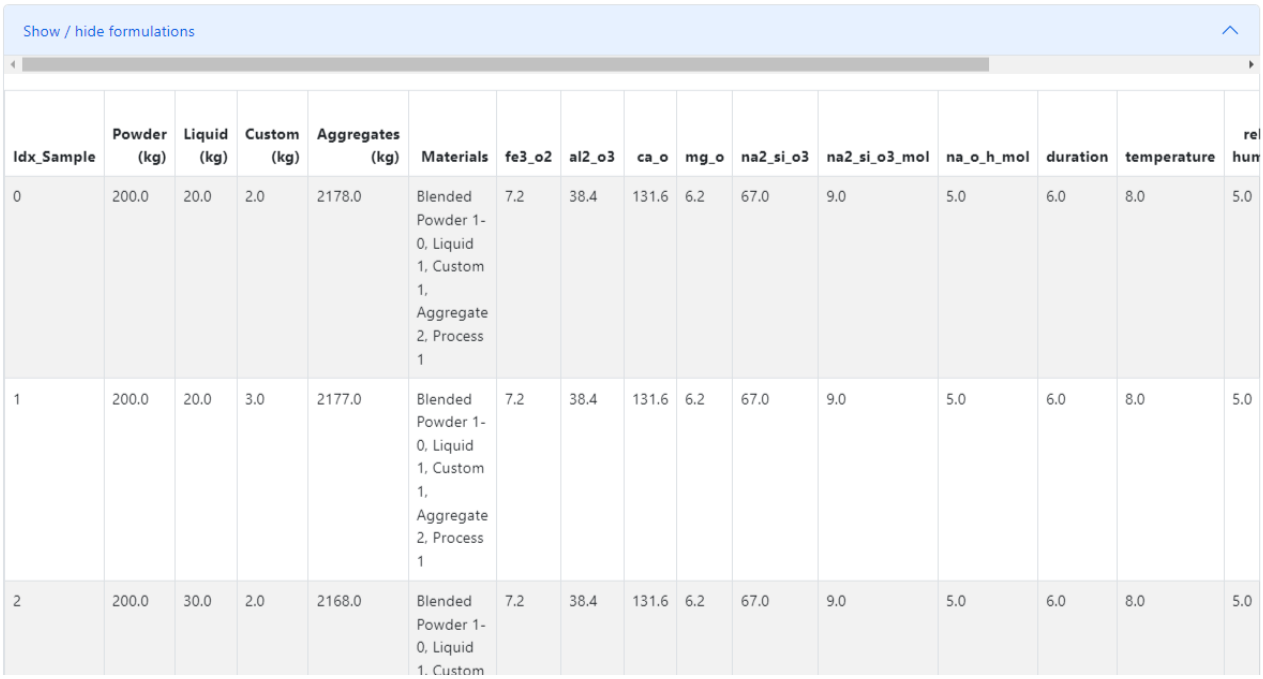
In addition to the selected materials and processes one needs to define a weight constraint. This corresponds to the total weight of the mix of materials to be used for the formulation. Further, one can optionally specify a name (this will, however, become relevant later).
Having chosen a valid configuration (currently it is required to have at least one powder, one liquid and one aggregate and a non-empty weight constraint), a form for specifying the various weight configurations can be requested.

This form works similar to the analogous form for blending materials (see description above) with the difference that the autocompletion features do not restrict the sum of each column to be 100%, but rather to be equal to the specified weight constraint.
Further, we also show the chosen processes for completeness. Having filled all fields on can proceed to the next step by clicking on the button at the bottom of the current screenshot. As a result, an overview of all weight combinations as determined by the specified
weight configuration is shown. Here the user can still decide to delete or edit some of the combinations (however, one needs to respect the pattern as described in the text in the screenshot below):



In the next step (clicking on the button to create material formulations for the given configuration), a formulation batch is created (the screenshot shows

only a couple of rows):



Idx_Sample	Powder (kg)	Liquid (kg)	Custom (kg)	Aggregates (kg)	Materials	fe3_o2	al2_o3	ca_o	mg_o	na2_si_o3	na2_si_o3_mol	na_o_h_mol	duration	temperature	rel hum
0	200.0	20.0	2.0	2178.0	Blended Powder 1-0, Liquid 1, Custom 1, Aggregate 2, Process 1	7.2	38.4	131.6	6.2	67.0	9.0	5.0	6.0	8.0	5.0
1	200.0	20.0	3.0	2177.0	Blended Powder 1-0, Liquid 1, Custom 1, Aggregate 2, Process 1	7.2	38.4	131.6	6.2	67.0	9.0	5.0	6.0	8.0	5.0
2	200.0	30.0	2.0	2168.0	Blended Powder 1-0, Liquid 1, Custom	7.2	38.4	131.6	6.2	67.0	9.0	5.0	6.0	8.0	5.0

For all the combinations of chosen materials and processes, a dedicated rows for all weight combinations are created. E.g. for the above configuration there are 8 weight configurations, 2 powders, 1liquid, 1 custom
and 1 process leading to a total of 8x2x1x1x1=16 rows.
One can now do the same steps over again to create another batch. Say e.g. the next batch would lead to 20 rows, the resulting formulation now consists of 36 rows. Once all batches have been added, the formulation can be
saved as a dataset by clicking the submit button (or delete using the delete button):

After submission, the dataset is saved and can be found when navigating to the materials discovery page.

4. Material Discovery

On the discovery page all your previously created datasets are shown. Further, you can also upload new datasets directly in terms of a csv upload (with utf-8 encoding).

Make sure you properly format your csv file. Allowed formats are

- column separator: "," and decimal delimiter "."
- column separator: ";" and decimal delimiter ","

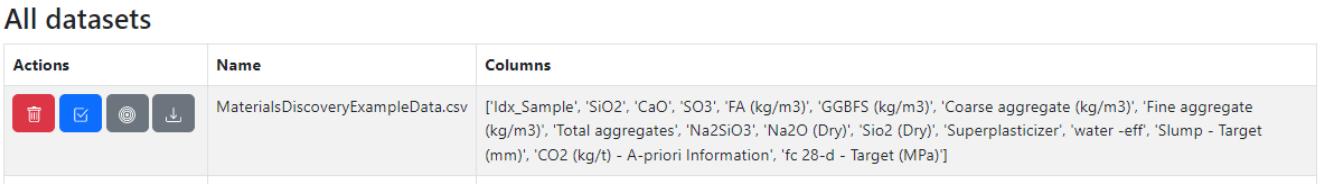
In the hypothetical case of a csv with two columns called col1 and col2 with only one row with values 1.2 and 2.1 a valid format would look like:

```
col1,col2
1.2,2.1

or

col1;col2
1,2;2,1
```

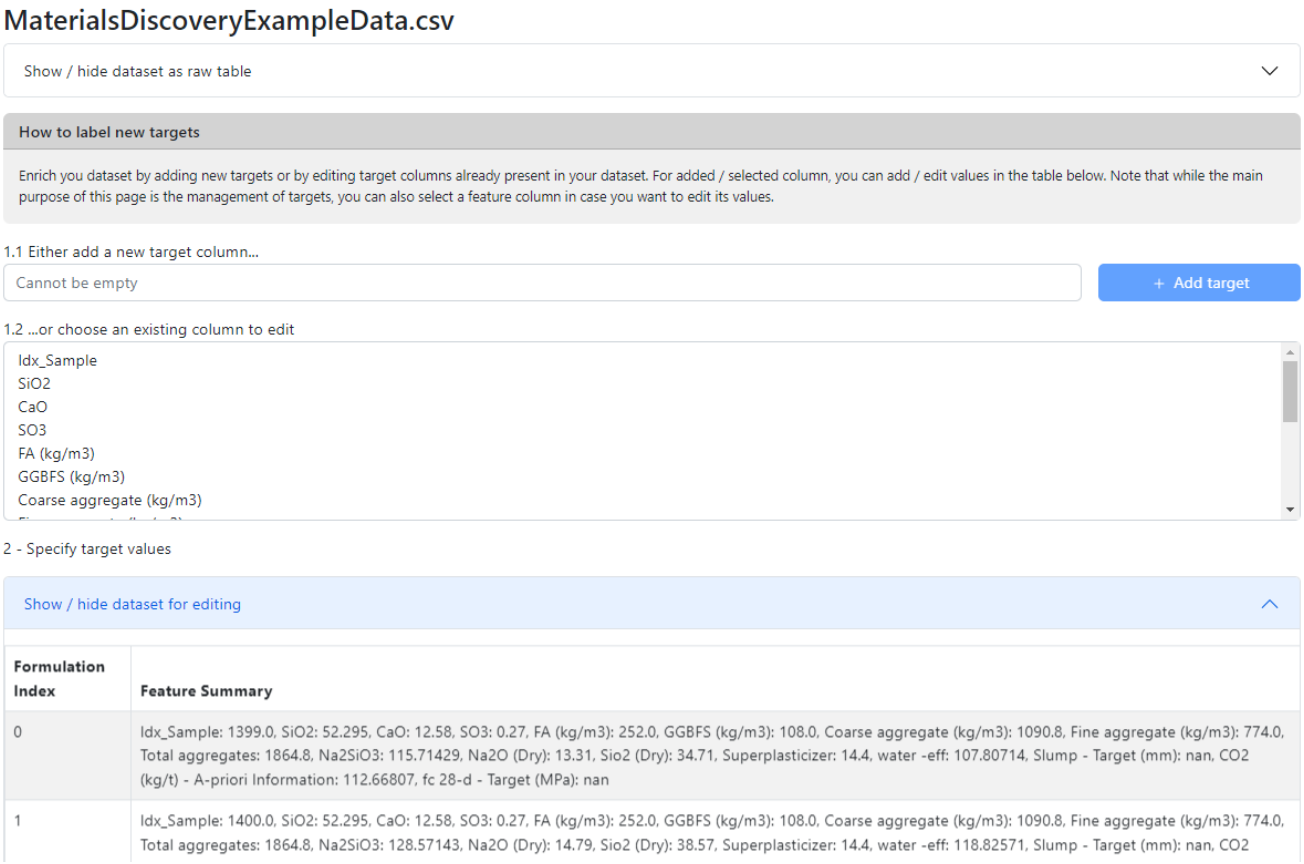
After successful upload the file is shown in a table:



A given dataset can be deleted, selected for running an experiment or extended by specifying target values.

4.1 Specify Targets

For specifying targets, press the 'bullseye'-button for the dataset of interest. You are shown a page similar to the one in the screenshot below:



You can also inspect the data as a dataframe by pressing the 'Show / Hide'-button. Further, it is possible to specify new target columns. First define a new column and after adding the target, you can set its values.
Note that you are not restricted to one target. Rather, you can add new target values one after the other. For example, in the next screenshot, two targets, X and Z were added and some known values
were filled and saved. Futher, there is the possibility to edit existing columns (in particular targets) by choosing from the selection field. This is relevant e.g. when an actual experiment after a prediction was performed
so that now the true label of a data row is known. Then, the value can be added and, using this enriched data, new predictions can be made thereby realizing the concept of sequential learning.

4.2 Configuring and runing an experiment

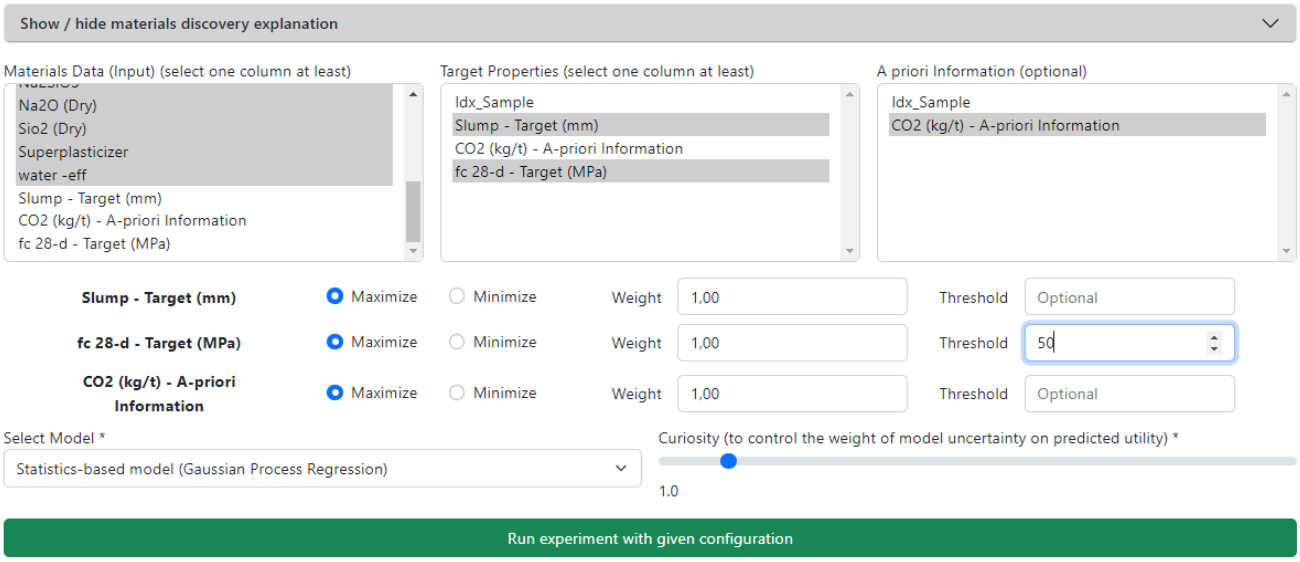
Having configured all targets, the set can be used for an experiment by selecting it in the discovery tab. You can select features, targets and a-priori information in the first step. Note that in principle all
columns are listed to be used as feature. Here, one should select only the ones which actually have the meaning of a feature. The columns which are not selected are automatically listed as suggestions for targets.
Once chosen, the remaining columns can be chosed as a-priori information. With the given selection, you can now define whether a given target / a-priori information should be minimized or maximized and how important
it is for your prediction (weight input). Furthermore you might specify a threshold for the selected targets / a-priori information. This makes sense in case you are only interested in finding e.g. a material with "50 Mpa strength or more"
without caring about the exact value. The algorithm processes it as follows:

- Minimizing a target with threshold T: Target predictions are not affected. However, all predicted target values smaller or equal than T equally contribute to the utility
- Maximizing a target with threshold T: Target predictions are not affected. However, all predicted target values bigger or equal than T equally contribute to the utility
- Minimizing an a-priori information with threshold T: All rows with values for the chosen a-priori information above T are simply dropped no utility for such rows is computes
- Maximizing an a-priori information with threshold T: All rows with values for the chosen a-priori information beelow T are simply dropped no utility for such rows is computes

Finally a model can be chosen and a curiosity value specified. The latter is a measure of how much (high curiosity) uncertainty we allow in our prediction. The higher the value, the more
we seek for new material candidates with properties which have not been explored so far.

Note that for the algorithm to work certain criteria must be met by the selection:

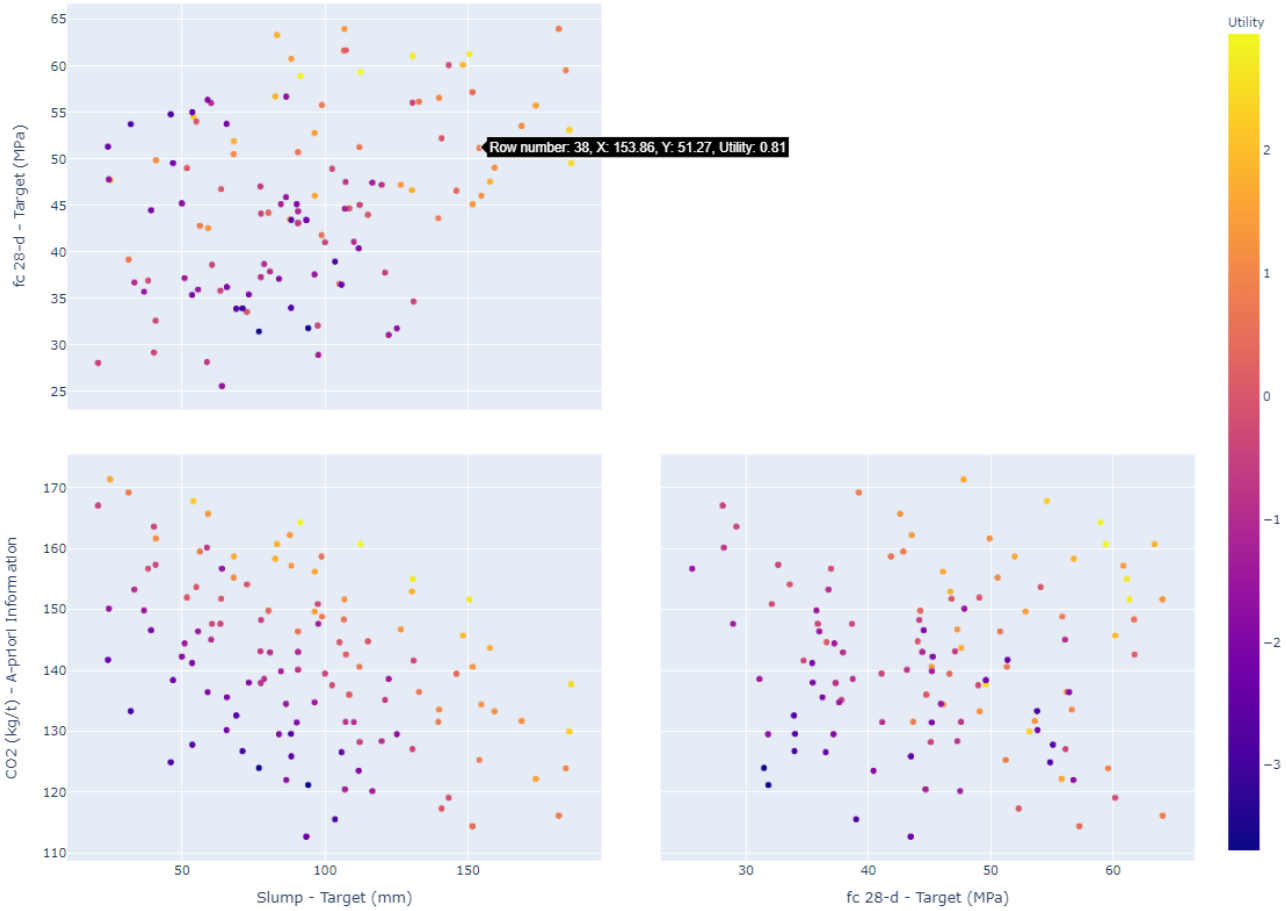
a) not all rows of the chosen targets are labelled
b) at least one of the selected target rows must be labelled
c) when choosing multiple targets for the Statistics-based model, make sure the same rows are labelled.
d) feature columns (Materials Data (Input)) containing at least one empty value are dropped by the algorithms
e) lolo Random Forest predicts slightly different result on each run for a given dataset and configuration (this is a known issue in the used package). The major conclusions for Sequential Learning are, however, unaffected.



Running the experiment by clicking on the button at the bottom of the screenshot, we get our predictions shown as a dataframe. It is sorted by utility such that the first row represents the most promising candidate
with respect to the parameters we have chosen in the screenshot above.

Row number	Utility	Novelty	Slump - Target (mm)	fc 28-d - Target (MPa)	Uncertainty (Slump - Target (mm))	Uncertainty (fc 28-d - Target (MPa))	Idx_Sample	SiO2	CaO	SO3	FA (kg/m3)	GGBFS (kg/m3)	Coarse aggregate (kg/m3)	Fine aggregate (kg/m3)	aggr
1	2.946420	0.523664	112.342442	59.452994	23.26243	7.68811	1430	47.085	18.30	0.45	210	210	966.00000	810.60000	1776
2	2.808785	0.620889	91.353715	59.014729	26.62892	9.16342	1588	47.085	18.30	0.45	215	215	967.96105	795.03895	1763
3	2.722163	0.262025	150.456857	61.376902	13.29414	3.70176	1429	47.085	18.30	0.45	210	210	966.00000	810.60000	1776
4	2.709731	0.401033	130.591258	61.171319	19.37737	5.30052	1558	47.085	18.30	0.45	215	215	967.96105	795.03895	1763

You can inspect the predictions more closely in terms of plots which are created in terms of the predicted targets and utility. Hovering over the point shows information regarding the given point.
Additionally, there is a button for downloading the prediction data. An excel is created containing sheets for the predicted data, the original data and the configuration that was used for making the prediction:


fc 28-d - Target (MPa)
Row number: 38, X: 153.86, Y: 51.27, Utility: 0.81
CO2 (kg/t) - A-priorI Inform ation
Slump - Target (mm)
fc 28-d - Target (MPa)
Utility