

Press Release
January 11th, 2016

Latest version of Qlucore Omics Explorer aids data exploration and biomarker discovery

Enhanced features with inbuilt clustering and advanced classifiers with one mouse click.

Qlucore, a world leader in the development of bioinformatics software, has today unveiled Qlucore Omics Explorer 3.2, the latest version of its advanced data analysis software.

Researchers will be able to undertake deeper data exploration and biomarker discovery using the additional functionalities in version 3.2 for clustering and classification.

Clustering is used to find subgroups among data samples and to see whether the samples naturally distribute themselves into distinct clusters. The new clustering functionality will enhance both the data exploration and statistical verification modes of the program. It will further assist the user to find clusters and subgroups in data, and complements existing functionalities such as projection score, variance filter and PCA plots.

The second significant addition is the classification functionality, or predictive modeling. The user will be able to build advanced classifiers at the click of a button, selecting from a range of different models, and will also be able to classify data based on generated models. The classification functionality can be used as an alternative way to finding the most relevant variables (features) to explain a condition, and be used for classification of samples in a diagnostic context.

In addition, several new plot options are included, the bar plot, with a wide range of configurations and the Kaplan-Meier version of the line plot. Together, these new functionalities are well integrated, easy to access, and will enable researchers to get deeper data insights into their work.

The combination of the existing strong visualization and analysis methods within Qlucore Omics Explorer, and the new clustering and classification functionalities, will strengthen the unique offering for visual, comprehensive, user friendly and fast data analysis of complex and demanding problems.

Additional information about the new functionalities:

Clustering

The support of data exploration and identification of potential subgroups and clusters is extensive in Qlucore Omics Explorer. The combination of variance filtering controlled with projection score and visualization using dynamic PCA gives excellent possibilities to identify subgroups and clusters. With the new inbuilt k-means clustering it is possible to get unbiased cluster proposals to further improve the functionality and to split the samples into a pre-defined number of groups. In the heatmap plot it is possible to cluster data using hierarchical clustering (dendrogram) which also is a form of clustering. The choice of starting points impacts the quality of the clusters significantly and with the use of k-means++ the starting point is selected very carefully.

In determining the number of clusters in a data set and how the samples should be distributed across the clusters requires a step-wise approach. The silhouette plot includes an overview of how data is clustered and an indication on how well a sample fits into a cluster.

Classification

The classification functionality is divided into two distinct areas which are reached from two different tabs in the user interface. Building and using a classifier belong to the area called supervised methods, as compared to clustering which is unsupervised.

- Build Classifier
- Classify

The Build Classifier offers easy access to advanced functionality for creation of a classifier based on the active data set. The objective is to, based on the active data set and a given sample annotation (the Key), create a model that can predict which group a new sample shall belong to. There are three different types of classifiers included; kNN, Support Vector Machines(SVM) and Random Trees. Working with classifiers it is central to validate the performance of the classifier on data not used in the building of the classifier. If a classifier is too specific towards the data used for creation the terminology overfitting is often used. There are two built in methods to perform validation:

- Using an external data set
- Using a cross validation scheme on the active data set, where the data set is divided into smaller training and test sets multiple times

The program will perform validation automatically.

When the classifier is built it can be used to classify new samples. The variables used in the classifier must be present in the data set to classify. The output from the Classify function is a new sample annotation describing how the sample(s) was classified.

New Plots and added configurability

The bar plot is a completely new plot type and the silhouette plot is a special case of this plot. The bar plot allows extensive configurations on what to include. The line plot has been extended with ROC/AUC configurations and it is also possible to create Kaplan-Meier survival plots based on an arbitrary annotation including survival time.

END

Qlucore started as a collaborative research project at Lund University, Sweden, supported by researchers at the Departments of Mathematics and Clinical Genetics, in order to address the vast amount of high-dimensional data generated with microarray gene expression analysis. As a result, it was recognised that an interactive scientific software tool was needed to conceptualise the ideas evolving from the research collaboration.

The basic concept behind the software is to provide a tool that can take full advantage of the most powerful pattern recogniser that exists - the human brain. The result is a core software engine that lets the user handle and filter data and the same time instantly visualise it in 3D. This will aid the user in identifying hidden structures and patterns. Over the last four years major efforts have been made to optimise the early ideas and to develop a core software engine that is extremely fast, allowing the user to explore and analyse high-dimensional data sets with the use of a normal PC, interactively and in real time.

Qlucore was founded in early 2007 and have clients in more than 25 countries. Qlucore Omics Explorer, represents a major step forward for easy data analysis with biologists as the main user group. Fast and interactive visualizations are combined with advanced statistics support, streamlined workflows for multiple data types, and a wide selection of presentation methods to aid the user. The presentation methods range from an innovative use of principal component analysis (PCA) to

interactive heat maps and flexible scatter plots. All user action is at most two mouse clicks away. The company's customers are mainly from the Life-science and Biotech industries.

Company Contact:

Carl-Johan Ivarsson

Phone: +46 46 286 3110

Email: carl-johan.ivarsson@glucore.com

Web: www.glucore.com

Press Contact:

Alison Scarrott, CBC

Tel: +44 (0)1483 537 890

Email: alisons@chazb.com