

## Quality assessment for intrinsic disorder predictions

Lukasz Kurgan

Department of Computer Science, Virginia Commonwealth University, Richmond, USA

### Abstract:

Intrinsic disorder prediction is being pursued for over 30 years. Modern predictors rely heavily on machine learning and are widely used in computational and experimental studies. However, they lack quality assessment (QA) scores that quantify which residue-level predictions are more likely to be correct. The complicating factor is that the QA scores must be optimized for specific disorder predictors since these methods rely on different types of disorder annotations and have substantially different predictive architectures.

We will discuss first-of-its-kind toolbox of methods that provide accurate QA scores for ten popular disorder predictors. The QUARTER (QUality Assessment for pRotein inTrinsic disordEr pRedictions) tool relies on a machine learning model that is optimized for each of the ten disorder predictors, guided by an empirical feature selection and disorder predictor-specific putative propensities for disorder. Empirical tests on a large test dataset reveal that QUARTER generates high quality scores which significantly outperform the disorder propensities output by the original predictors. QUARTER is available as a convenient webserver at <http://biomine.cs.vcu.edu/servers/QUARTER/>.

The main application of the QA scores produced by QUARTER is annotation/selection of a high-quality subset of residue-level disorder predictions. We show that when combining results from the ten disorder predictors, the QA scores can be used to identify 40% of residues for which disorder is predicted with 95% precision.