Analysis of LIBS amino acid spectra and the impact of neighborhood size on the efficacy of nonlinear analysis

Piyush Kumar Sharma¹, Gary Holness², and Poopalasingam Sivakumar, Yuri Markushin, Nouredinne Melikechi³

¹Department of Mathematical Sciences, ²Department of Computer and Information Sciences, ³Department of Physics and Engineering

Delaware State University, Dover, DE 19901

Abstract

In recent work we explored the impact of LIBS spectra on suitable neighborhood size over which to consider pattern phenomena, if nonlinear methods capture pattern phenomena with increased efficacy, and how they improve classification and identification of compounds. We analyzed four amino acids, polysaccharide, and a control group, water. We developed an information theoretic method for measurement of Laser Induced Breakdown Spectroscopy (LIBS) energy spectra, implemented manifold methods for nonlinear dimensionality reduction, and found while clustering results were not statistically significantly different, nonlinear methods lead to increased classification accuracy. The goal of the poster is to discuss recently published results.

Introduction

For a given D dimensional dataset point find low dimensional, d(<< D), embedding Y by algorithm:[s].

LIBS Experimental Setup.

LLE

• Find K nearest nbhd for each point x_i ∈ X.
• Compute a set of weights w_i for each point that best describe the point as a linear combination of its neighbors. Solve for reconstruction weights w_i.
• Find embedding Y by using an eigenvector based optimization technique which minimizes reconstruction loss.

Clustering

Davies-Bouldin Index = 1/k ∑_i=1^K max(D_i)/d_i, where D_i = ∑_{j≠i}||x_i - x_j||, where d_i is the within-to-between cluster distance ratio for the i_th and j_th clusters, d_i is the average distance between each point in the i_th cluster and the centroid of the i_th cluster, d_j is the average distance between each point in the i_th cluster and the centroid of the j_th cluster. d_i is the Euclidean distance between the centroids of the i_th and j_th clusters.

Classification

In support vector machine (SVM), decision surface separating the classes is a hyperplane w^T x + b = 0 where w is a weight vector, x is an input vector, and b is a bias. The vectors that define the hyperplane are the support vectors. The problem can be expressed in dual form with w = ∑_i a_i φ(x_i): min_{a_i} D(a) = ∑_i a_i φ(x_i) φ(x_j) - ∑_i a_i such that ∑_i a_i ≤ C

Conclusion

1. LIBS pattern phenomena are nonlinear.
2. k-neighborhood size:
   a. for inter sample variabilities is [8; 15]
   b. for intra sample variabilities is k > 15
3. Identified micro-wells influence on D-Serine spectra.
4. Future work will comprise remaining amino acids.

Acknowledgement

This work was supported by NSF Award#: HRD-1242067, Award#: CNS-1205426, and Propinquo LLC

Bibliography