

Predicting Pectin Performance Strength using Near Infrared Spectroscopic Data

A Comparative Evaluation of 1d-CNN, PLS and Ridge Regression Modeling

Kasper Einarson¹ | Andreas Baum¹ | Terkel B.Olsen¹ | Jan Larsen² | Ibrahim Armagan² | Paloma A.Santacoloma² | Line K.H. Clemmensen¹

¹ Department of Applied Mathematics and Computer Science, DTU, Kgs Lyngby, Denmark

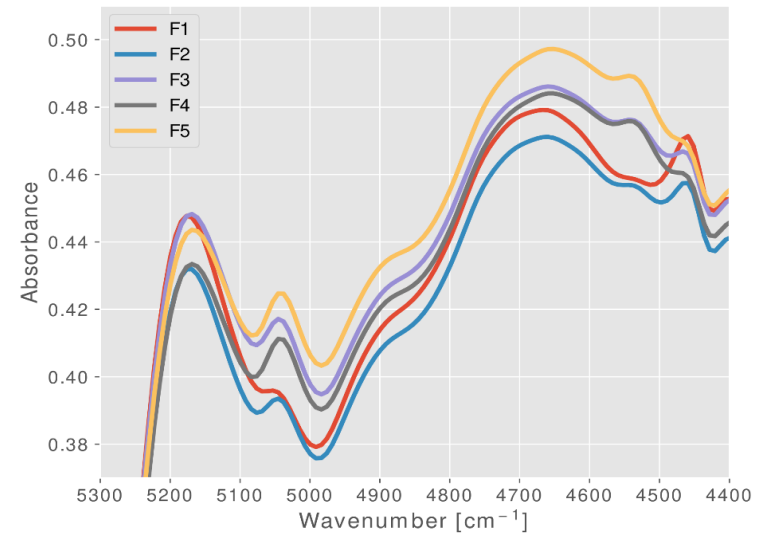
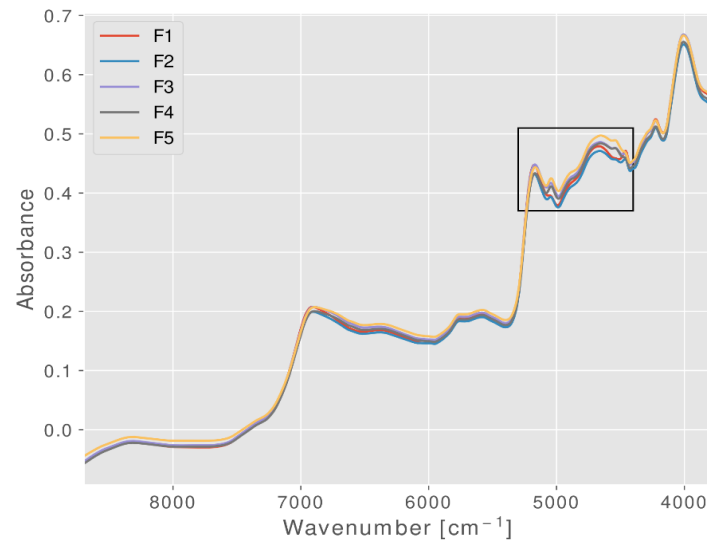
² CP Kelco ApS, Denmark

Goal of the project

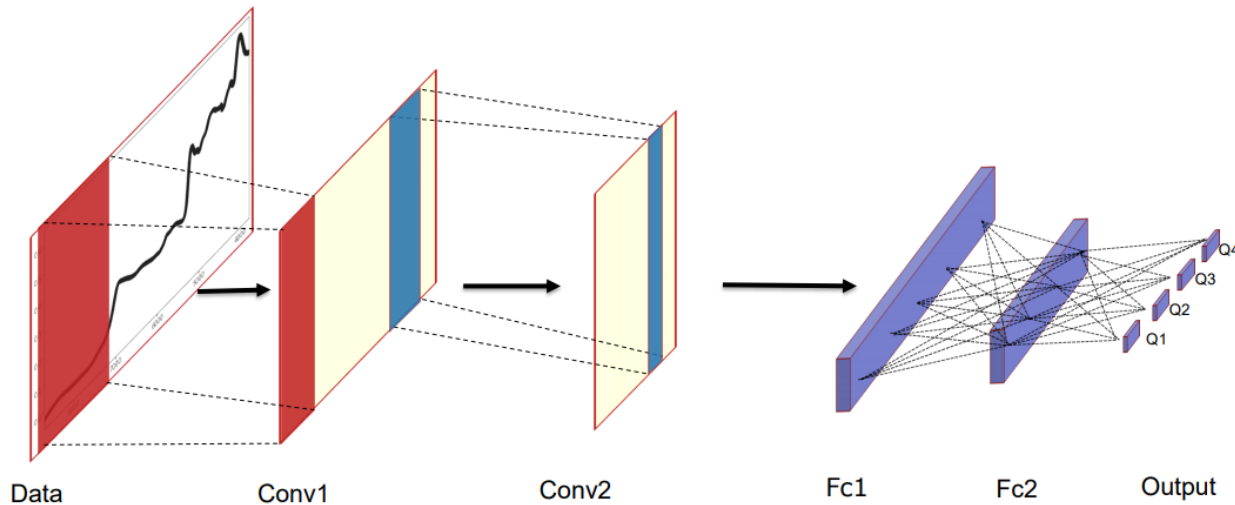
To predict the physical properties of five different industrial pectin formulations based on their near infrared spectral data.

Data

- NIR spectra on 888 pectin samples across five different groups/formulations (F1 – F5)
- 4 Gel Performance Strengths are provided for each spectrum (multivariate response)

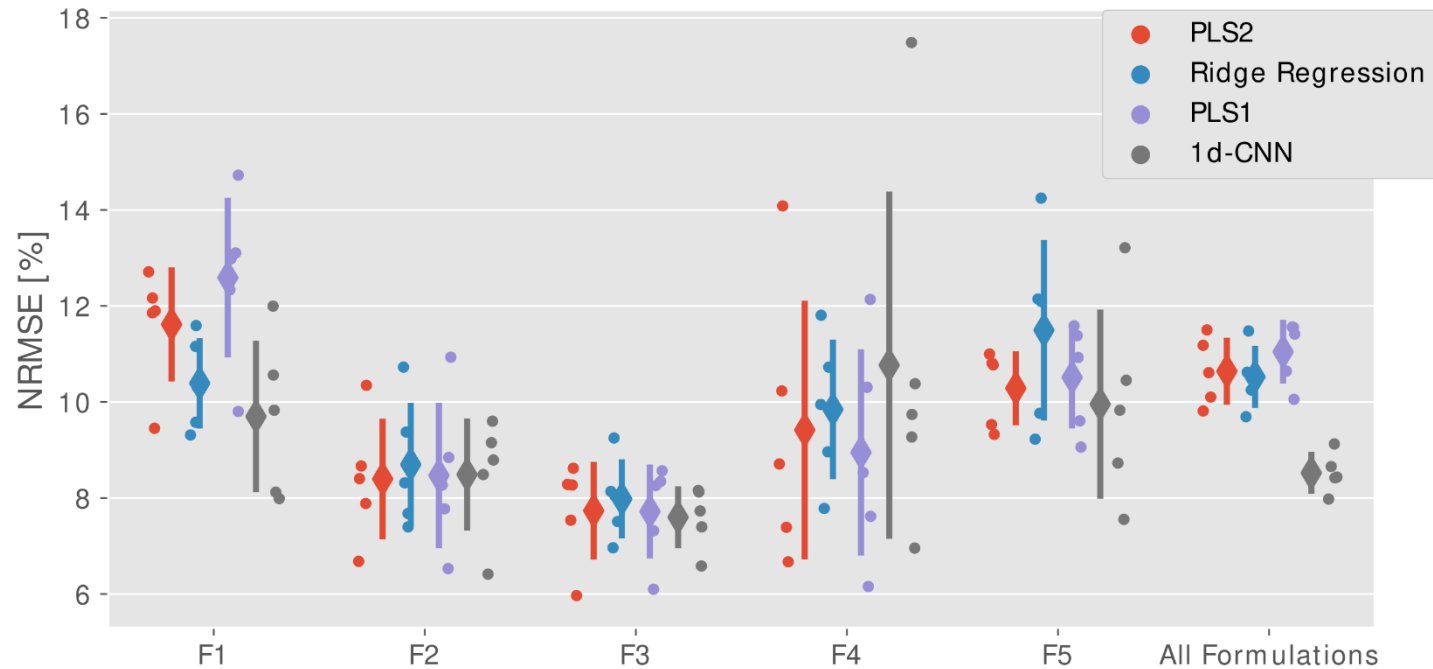


Methods



- Pre-processing: SNV + Autoscaling
- Model Performance compared for 4 Methods:
 - PLS1
 - PLS2
 - Ridge Regression
 - 1-D Convolutional Neural Network (1d CNN)
- Nested Cross Validation
 - Hyperparameter Tuning
 - Estimation of Generalization Error

Results



- Equal performance within groups (F1 – F5)
- 1d-CNN indicates lower NRMSE when modeling all groups jointly

DTU



Thank you for your attention!