Case-Study 15

Amyotrophic Lateral Sclerosis (ALS)

Overview: This case-study examines the patterns, symmetries, associations and causality in a rare but devastating disease, amyotrophic lateral sclerosis (ALS). ALS demands conducting clinical trials and collecting big, multi-source and heterogeneous datasets that can be interrogated to derive potential biomarkers. Overcoming many scientific, technical and infrastructure barriers is required to establish complete, efficient, and reproducible protocols (pipelines/workflows) starting with acquiring raw data, preprocessing, aggregation, harmonization, analysis, visualization and result interpretation.

The clinical data shows that the rate of ALS progression varies significantly among patients. Majority of the patients die within 3 to 5 years after ALS onset, however, a few are able survive for over 10 years. This heterogeneity of disease course hinders demonstration of its biological mechanism and development of effective treatment. We need to develop reliable predictive models of ALS progression to understand the pathophysiology of the disease.

Driving Challenges:
- o   What patient phenotypes can be automatically and reliably determined?
- o   Predict the change of the ALSFRS slope change using the holistic patient-specific data.
- o   Predict survival of patients at a given time-point (post diagnosis).

Meta-Data

- There are 2 datasets:
    - o   training (N1=2,223): **ALS_TrainingData_2223.csv**, and
    - o   testing (N2=78): **ALS_TestingData_78.csv**

- Each dataset includes the following 131 variables:

    ID; Age_mean; Albumin_max; Albumin_median; Albumin_min; Albumin_range; ALSFRS_slope; ALSFRS_Total_max; ALSFRS_Total_median; ALSFRS_Total_min; ALSFRS_Total_range; ALT.SGPT._max; ALT.SGPT._median; ALT.SGPT._min; ALT.SGPT._range; AST.SGOT._max; AST.SGOT._median; AST.SGOT._min; AST.SGOT._range; Basophils_max; Basophils_median; Basophils_min; Basophils_range; Bicarbonate_max; Bicarbonate_median; Bicarbonate_min; Bicarbonate_range; Bilirubin..total._max; Bilirubin..total._median; Bilirubin..total._min; Bilirubin..total._range; Blood.Urea.Nitrogen..BUN._max; Blood.Urea.Nitrogen..BUN._median; Blood.Urea.Nitrogen..BUN._min; Blood.Urea.Nitrogen..BUN._range; BMI_max; bp_diastolic_max; bp_diastolic_median; bp_diastolic_min; bp_diastolic_range; bp_systolic_max; bp_systolic_median; bp_systolic_min; bp_systolic_range; Calcium_max; Calcium_median; Calcium_min; Calcium_range; Chloride_max; Chloride_median; Chloride_min; Chloride_range; Creatinine_max; Creatinine_median; Creatinine_min; Creatinine_range; Eosinophils_max; Eosinophils_median; Eosinophils_min; Eosinophils_range; Gender_mean; Glucose_max; Glucose_median; Glucose_min; Glucose_range; hands_max; hands_median; hands_min; hands_range; Hematocrit_max; Hematocrit_median; Hematocrit_min; Hematocrit_range; Hemoglobin_max; Hemoglobin_median; Hemoglobin_min; Hemoglobin_range; leg_max; leg_median; leg_min; leg_range; Lymphocytes_max; Lymphocytes_median; Lymphocytes_min; Lymphocytes_range; Monocytes_max; Monocytes_median; Monocytes_min; Monocytes_range; mouth_max; mouth_median; mouth_min; mouth_range; onset_delta_mean; onset_site_mean; Platelets_max; Platelets_median; Platelets_min; Potassium_max; Potassium_median; Potassium_min; Potassium_range; pulse_max; pulse_median; pulse_min; pulse_range; Red.Blood.Cells..RBC._max; Red.Blood.Cells..RBC._median; Red.Blood.Cells..RBC._min; Red.Blood.Cells..RBC._range; respiratory_max; respiratory_median; respiratory_min; respiratory_range; Sodium_max; Sodium_median; Sodium_min; Sodium_range; SubjectID; trunk_max; trunk_median; trunk_min; trunk_range; Urine.Ph_max; Urine.Ph_median; Urine.Ph_min; Urine.Ph_range;

White.Blood.Cell..WBC._max; White.Blood.Cell..WBC._median; White.Blood.Cell..WBC._min; White.Blood.Cell..WBC._range