Thousands of algorithms trained for predicting the treatment efficacy of rheumatoid arthritis

24 August 2016

Rheumatoid arthritis is a chronic inflammatory autoimmune disorder affecting millions of people worldwide. Anti-TNF treatment is a widely used treatment blocking the inflammatory cytokine, but it fails in approximately 1/3 of the patients.

The objective of the wide crowdsourced study was to use algorithms in assessing the efficacy of anti-TNF treatment based on clinic and genetic data, or in identifying the non-responders before the treatment. 73 research teams, altogether hundreds of researchers, worldwide competed in an open challenge using the most comprehensive data available of more than 2700 patients and using a wide range of state-of-the-art modeling methodologies.

The eight teams with the best predictive performances were invited to the final phase. Team MI of Aalto University and Helsinki University Institute for Molecular Medicine Finland (FIMM) were among those eight teams.

"We used both sparse linear regression model and multiple kernel learning model to predict the treatment response based on the genetic and clinical information, describes Lu Cheng, ‡ Postdoctoral Researcher at the Department of Computer Science.

Team Outlier, the winner in the final phase, did not use any genetic information in the final round. As a conclusion the currently collected genetic data did not significantly contribute to the prediction of treatment response above the clinical predictors including sex, age and medical information.

"If a limited amount of genetic variants would explain the failure of the treatment in some of the patients, we would have had the prediction model as a result of a vast study like this. Either the amount of the genetic variants is much bigger and their effects respectively much smaller, or the missing heritability is better explained by genetic variants not included in the study, such as rare variants," tells University Lecturer Pekka Marttinen.

Over the course of the 16-week algorithm training period, 73 teams submitted a total of 4874 predictions for evaluation. The research results have been published in Nature Communications.


Leaderboard of the crowdsourced research challenge initial phase, Team MI ranking third.

Rheumatoid arthritis is a chronic inflammatory autoimmune disorder affecting millions of people worldwide. Anti-TNF treatment is a widely used treatment blocking the inflammatory cytokine, but it fails in approximately 1/3 of the patients.

The objective of the wide crowdsourced study was to use algorithms in assessing the efficacy of anti-TNF treatment based on clinic and genetic data, or in identifying the non-responders before the treatment. 73 research teams, altogether hundreds of researchers, worldwide competed in an open challenge using the most comprehensive data available of more than 2700 patients and using a wide range of state-of-the-art modeling methodologies.

The eight teams with the best predictive performances were invited to the final phase.

Team MI of Aalto University and Helsinki University Institute for Molecular Medicine Finland (FIMM) were among those eight teams.

"We used both sparse linear regression model and multiple kernel learning model to predict the treatment response based on the genetic and clinical information, describes Lu Cheng, ‡ Postdoctoral Researcher at the Department of Computer Science.

Team Outlier, the winner in the final phase, did not use any genetic information in the final round. As a conclusion the currently collected genetic data did not significantly contribute to the prediction of treatment response above the clinical predictors including sex, age and medical information.

"If a limited amount of genetic variants would explain the failure of the treatment in some of the patients, we would have had the prediction model as a result of a vast study like this. Either the amount of the genetic variants is much bigger and their effects respectively much smaller, or the missing heritability is better explained by genetic variants not included in the study, such as rare variants," tells University Lecturer Pekka Marttinen.

Over the course of the 16-week algorithm training period, 73 teams submitted a total of 4874 predictions for evaluation. The research results have been published in Nature Communications.
