

Klassificering av ALMA-data med hjälp av maskininlärning

Bakgrund

Atacama Large Millimeter/submillimeter Array (ALMA) är ett observatorium beläget på 5000 m höjd över havet i norra Chile. Det består av 66 antenner som verkar tillsammans på ett sådant sätt att mycket högupplösta bilder kan tas. De objekt som ALMA observerar sträcker sig från vårt eget solsystem till avlägsna galaxer.

Problembeskrivning

Det publika ALMA-arkivet innehåller idag data från nästan 60 000 unika observationer. För astronomer innebär detta ofta ett tidskrävande arbete. Manuell inspektion av hela arkivet är helt enkelt inte möjligt. Ofta är det så att en speciellt intressant observation är unik i sitt slag. Det är då intressant att ta reda på om det tidigare genomförts observationer med liknande resultat. Ett sådant exempel är sökandet efter stoft i vindar från protoplanetära skivor. Vi vet idag inte hur mycket material och vilket material som lämnar protoplanetära skivor via vindar. Om det är så att stoft förs bort, vill vi veta hur mycket, och vilken storlek stoftkornen har. Det är just från stoft som planetbildningsprocessen startar, då små mikrometer-stora korn börjar klumpa ihop sig till större enheter. Om en del av det stoftet lämnar systemet, ja, då tror vi också att planetbildningen påverkas av det.

I observationer syns stoftet i vinden genom en avvikelse från en normalfördelning (signaturen av en protoplanetär skiva). Hittills har dock endast en handfull sådana bilder hittats. Vi vill därför veta om det finns fler unga solsystem som uppvisar denna karaktäristik.

Arbetsätt

I det här kandidatarbetet kommer ni använda er av programspråket *Python* för att konstruera ett beslutsträd som väljer ut bilder för vidare klassificering och analys. Med hjälp av redan kända observationer kommer träningsdata konstrueras. Denna träningsdata kommer sedan användas för att träna ett neuralt nätverk (CNN) för vidare klassificering av observationer. Viss bildbehandling i mellansteget mellan beslutsträdet och klassificeringen kommer behöva ske. Arbetet kommer främst vara till hjälp för automatisk klassificering av unga solsystem som har observerats med ALMA. Algoritmen kommer dock utvecklas på ett sådant sätt att den även kan komma till användning inom andra forskningsområden (t ex vid studier av galaxer, och gamla stjärnor) eller vid observationer med andra teleskop (t ex Square Kilometer Array).

Projektet kommer delvis att genomföras i samarbete med kollegor från University of Virginia och National Radio Astronomy Observatory (NRAO).

Målgrupp:

Teknisk fysik (TKTFY), Teknisk matematik (TKTEM), Kemiteknik med fysik (TKKEF), Elektroteknik (TKELT), Globala system (TKGBS), Medicinteknik (TKMED) samt Datateknik (TKDAT). Stort intresse för programmering krävs och en introduktionskurs i Data Science och AI är önskvärt.Handledning kommer att ske på svenska och engelska.

Gruppstorlek:

3–6 studenter, max 1 grupp

Handledare:

Per Bjerkeli, per.bjerkeli@chalmers.se

Maria Carmen Toribio, toribio@chalmers.se

Examinator:

Magnus Thomasson, magnus.thomasson@chalmers.se