Week 4 meeting notes

|  |
| --- |
| **Mass Loss function:**  poorly understood, poorly handled by MESA  The first epoch of mass loss occurs at the tip of the RGB  Mass loss happens up the RGB through pulsations or gas is stolen by companions  Then red clump stars have to shed mass to be able to form white dwarfs  Ignore for now but we could consider looking into it at some point  We can ignore this by only considering stars that are main sequence sub-giants or 1st ascent RGB  Because these likely have not undergone any significant mass loss (some may have done due to companions).  So restrict the stars we look at to no more than 8R☉. All red clump above this radius (subject to some metallicity arguments)  If you know the temp and distance to a star you can roughly estimate its radius |
| The loss function for neural nets: use it to tell if the neural net is doing a good or a bad job  quantitative difference between what the neural net is currently outputting and what you want it to output, is used to calculate the loss function  Standard type:  Mean Squared Error: MSE = 1/N Σ(Ytrue - Y)2  Over weights things that are more wrong  Under weights things that are slightly wrong    Want to find the minimum      Can get rid of the constants i.e. ½ and C  Minimize the negative log likelihood (maximizing the probability given the data and the model)    If sigma is the same for all the data points you can ignore that as well, which leaves the MSE equation.  Therefore the MSE is equivalent to minimizing the negative log likelihood if all the data has the same uncertainty.  If you have noisy data (Gaussian noise) you should use the MSE  Mean Absolute Error: MAE = 1/N Σ|(Ytrue – Y)|    From the neural net the input and the output for the stellar evolution code doesn’t have any noise on it, so you should use the MAE i.e. MESA doesn’t have any uncertainty on the output. |
| Project:  Get a grid from someone (the guy 2 doors down)  Put in fundamentals, get observables. What is the chance of me getting this observable if my underlying model is correct and then you tune the underlying model until you get the best chance of getting the data  3 things on our plate:   1. Be able to hierarchically model 2. Either take ages from somewhere or do a neural net for it ourselves 3. Physical sciency understanding of open clusters (similar ages and similar composition). This is required to construct the hierarchical prior/population prior?, which you can then adjust the population prior and the parameters of each star and the neural net tells you the parameters of each star.   If we had a grid of stellar models, for each model there are some stellar properties  MESA and GYRE tell us how to move from those stellar properties to observables.  For a star with some properties what are the chances of getting the data we actually see:  We would create a model, map it through to observables and compare the mapped observables with the true observables, to see how good the mapped observables are.  We plan on swapping out this step using MESA and GYRE and replacing it with a neural net.  Improve the neural net by fiddling with the fundamentals.  But we will also be doing it hierarchically.  We will have a set of stars with a prior because they all belong to the same open cluster and so the age of each star should be consistent with some distribution that describes the timescale of formation in the cluster. Which gives us a ‘over the top constraint’ on what these fundamental properties could be.  And then the data (which is at the bottom) goes up and the prior goes down for simultaneous constraint.  Instead of having to move one thing at a time to check how this effects the results, with the HMC “it all changes at the same time” and the HMC take care of the optimization and gives back samples of the posterior probability distribution i.e. it summarises the values we are interested in (whether that is the age of individual stars or the prior of the population ages).  Neural net is pretrained on data from MESA and GYRE. Once we’re happy that it is doing a good job, we fix the neural net in a optimally trained state.  Then dump it into the HMC machinery and we’re good to go.  “the sciency bit for all this is to infer the hyperpriors for a cluster i.e. age and mass” – Hin Leung  Mass will be difficult due to selection effects but interesting, ignore it for now.  Mg/Fe = is good because it tells you about what epoch is occurring until there is a turnover  It would be interesting to see if we could predict age as a function of magnesium for a cluster HOWEVER this is very difficult due to the errors or the magnesium will be large compared to how much the magnesium is changing and the uncertainty on the ages will be large too.  We could try and predict the ages of these stars to say if a certain star is older than another.  AIM to get the age spread/distribution of a cluster  Further: turn those ages into something interesting  We could take some models of cluster formation and predict what that corresponds to in aspread of alpha values (alpha is a proxy for magnesium).  Say the age spread is that 95% of stars were formed within 50 Myrs, then you would expect the earliest ones to have some alpha.  From this we could say: “here is what we predict for the age spread, is it consistent with what we see chemically?” |
| Coherent oscillations (occur in hotter stars): the oscillations always look like a sine wave  Sun like oscillations = stochastic oscillations: stochastically excited and randomly damped  For coherent modes of oscillation you can estimate their frequencies to a very high precision  But a neural net cannot approximate a frequency output from a stellar model to that precision  Which is bad because you want more noise from the observation than from the neural net (numerical proceses)  For red giants the observations will be the limiter for the precision |
| Backwards modelling is bad because you cannot get a probability of an observation given the model.  Can’t do Bayesian stuff |
| Random forests:  Don’t have the property that give you the back propagation  Neural nets allow you to calculate how the loss functions changes if you change any of the weights or biases but you cannot do this with random forests.  Random forests mitigate overfitting but do still overfit  Prevent a neural net from overfitting by controlling the architecture  The more layers and the more neurons per layer the more flexible the neural net is as a function.  You will underfit if the architecture is too small.  So you will need to train multiple architectures  If the architecture is too large it can overfit  Papers don’t tend to describe their architecture because it’s kind of trial and error to get the best output and there is no proper method to it. |
|  |
| Commit to github:  Create a folder “Harry’s paper notes”  Add comments to why things are important |
| Introduction:  Open clusters, why are they important and how they link to stellar evolution  What is stellar evolution  What models we have right now  What is asteoseismology  Make a “daft hierarchical model”  Don’t need to say we’re using karas but instead that we’re planning to use a neural net  Title draft: “Hierarchical modelling of open clusters, for the timescales of evolution”  Mark criteria:  Demonstrate understanding: strengths of what we’ve chosen to do.  Demonstrate that our project plan is sensible |
| We could calculate grids ourselves, we will talk about that later  M67 is good because it’s solar age  We don’t need to restrict ourselves to asteroseismic clusters, however the age uncertainties will be worse than with asteroseismic data  We might want to run tests on the neural net without any mode frequencies at all to check things are sensible, to make sure there aren’t any glaringly obvious errors.  Don’t run everything in one, check each step to make debugging easier.  You could email the first author of a paper if you need further results from a paper  “Hi I’m a year 4 student working with Dr Guy Davies, I’m looking to do BLAH can you help me with BLAH by providing me with BLAH”  Ask in small steps to get them invested, and then ask for more. |
| Wednesday 30th, 13:30 |
| Give quick context about general stuff  Asteroseismology allows us to study the interior of stars using pulsation frequencies and comparing models. |
| Create an issue on github and assign it to Guy  Give info about the issue and what file the issue is in, or whatever. |
|  |

Filled out Annual review forms