Diagnostics of Spectral Feature in Mg II line from the Chromoshpere

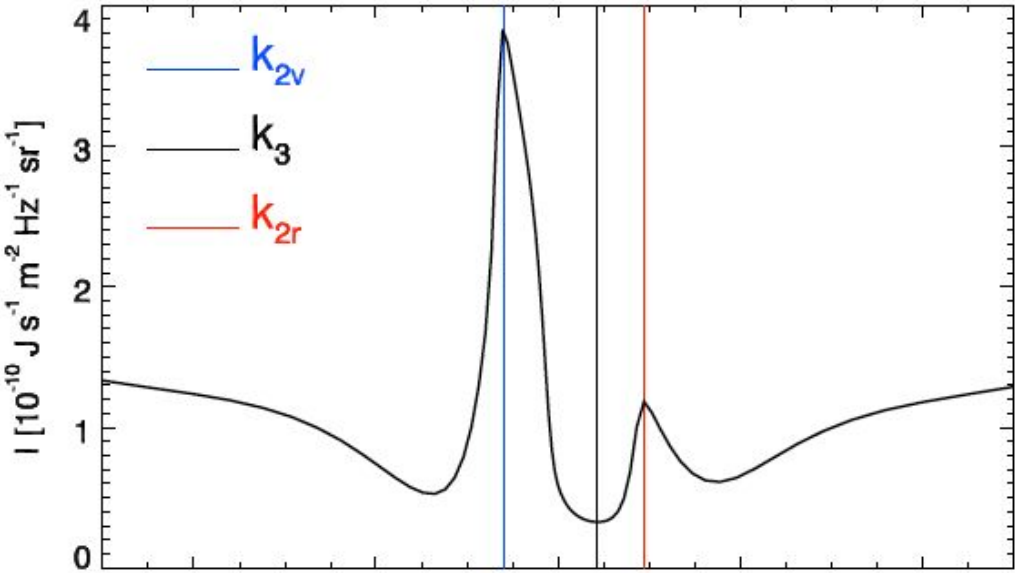
~Sambit

Overview

- Data
- Aim of project
- Features and Labels
- Correlation Test
- Regression
- Results
- Conclusion

DATA

- The data contains the chromospheric MgII line
- There are two special spectral features in the line, known as k & h feature
- **K-feature: (279.635 nm)**
  - K3 is a trough sandwiched between two asymmetric peaks called k2v and k2r
  - They are very sensitive to sun's atmosphere
- **h-feature:(280.353 nm)**
  - This feature is similar to k-feature



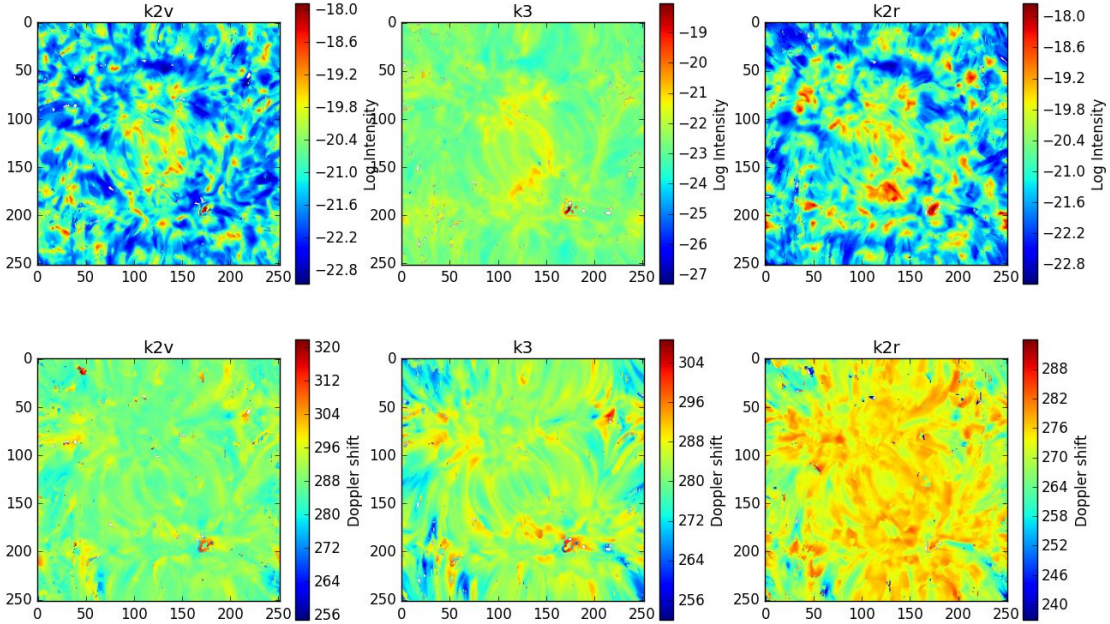
(Leenaarts et. al. 2013)

Aim of the project

- To predict the temperature of the solar atmosphere at an average height (T_avg) with minimum possible inputs from MgII line
- The h & k features have the maximum interaction with the atmosphere. Therefore, they contain the maximum information

Features and labels

- The intensity of the k & h features along with their doppler shifts are taken to be the features for the machine learning
- The plot shows the 252x252 images that are observed
- Each pixel corresponds to a different line profile. Therefore, there are 63504 data points

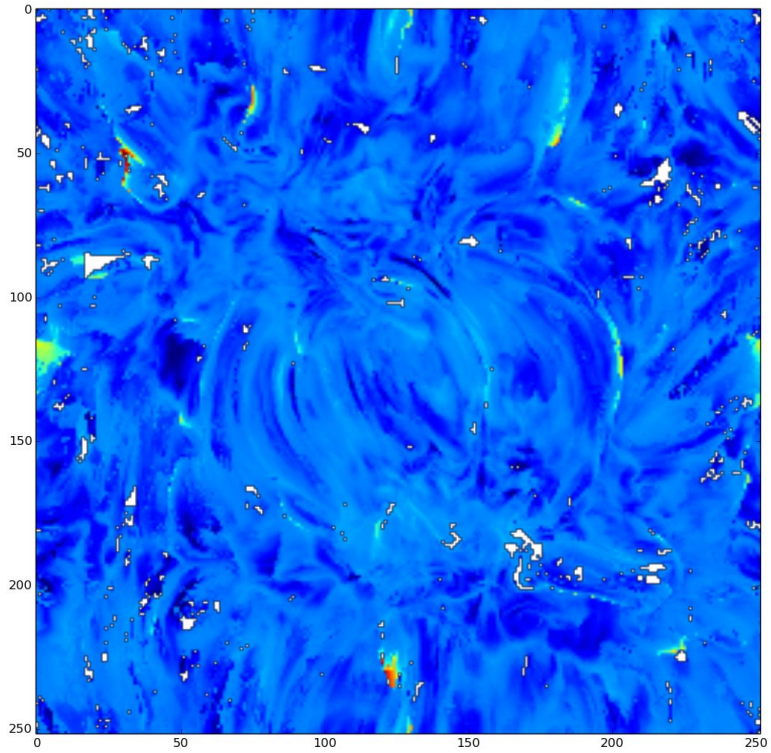


Features and labels II

- Apart from the intensities and the doppler shifts, the asymmetry of these values also contain information (Leenarts et. al. 2013)
- Thus, the asymmetries in the line parameters is also included as features using the following formula,

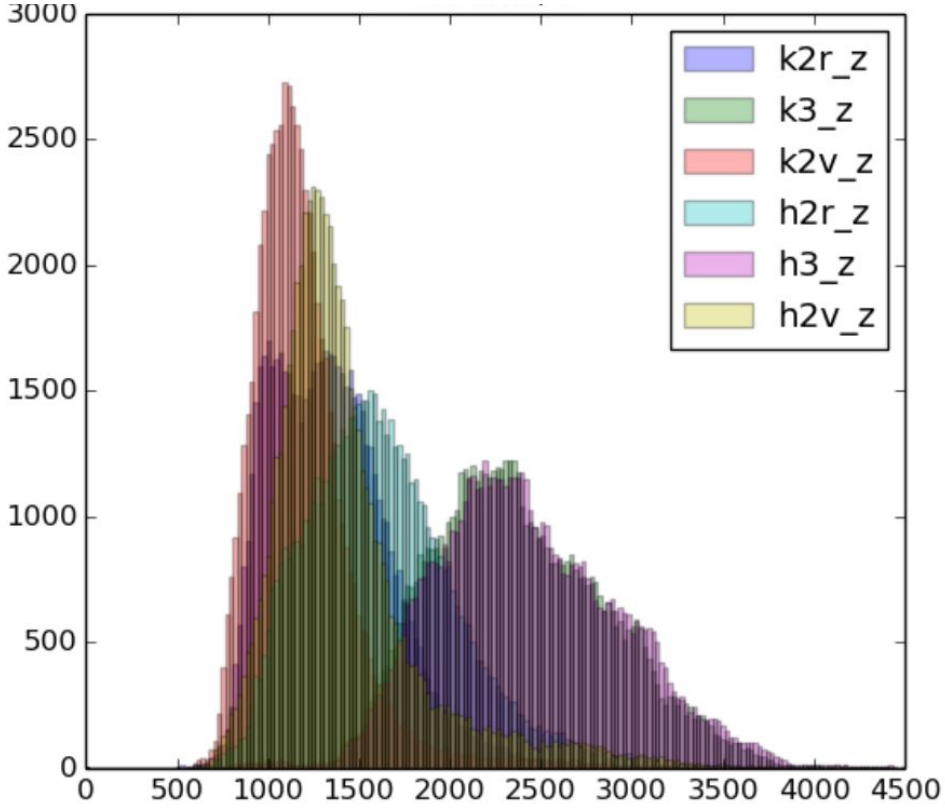
$$R = (A-B)/(A+B)$$

- The labels for the machine learning algorithm is the T_avg (shown in figure)
- The outputs from the MHD simulations is used for training the method



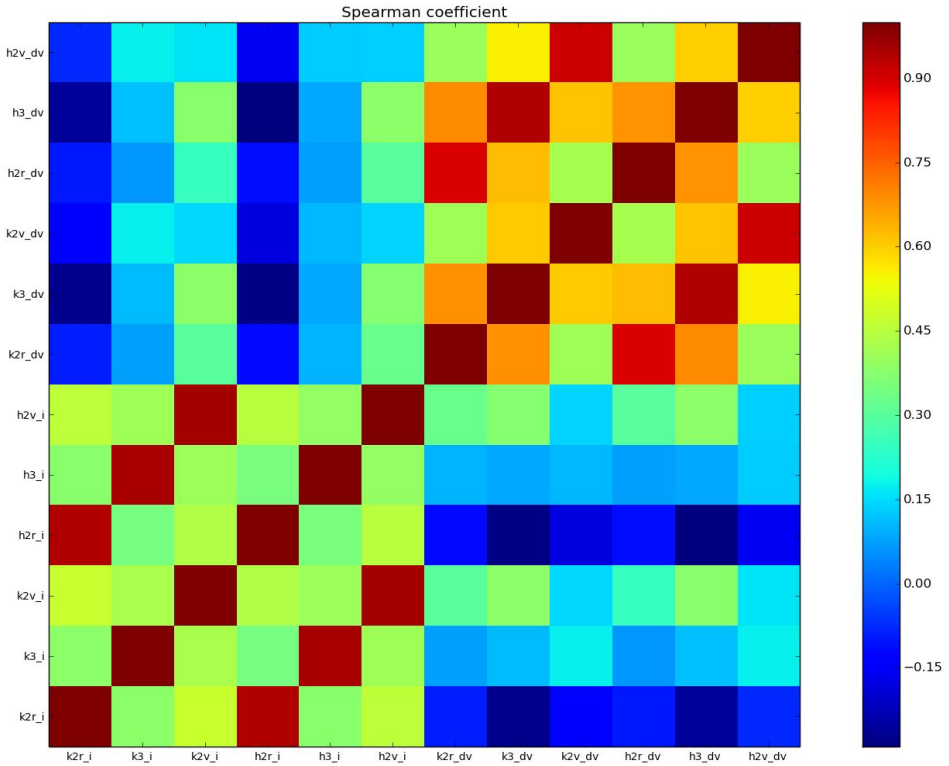
Response function (Labels)

- The point in the line interacts with the atmosphere at different height
- The response of the line features has a Gaussian-like shape
- From the figure we can see that there is a common height where they all interact with the atmosphere



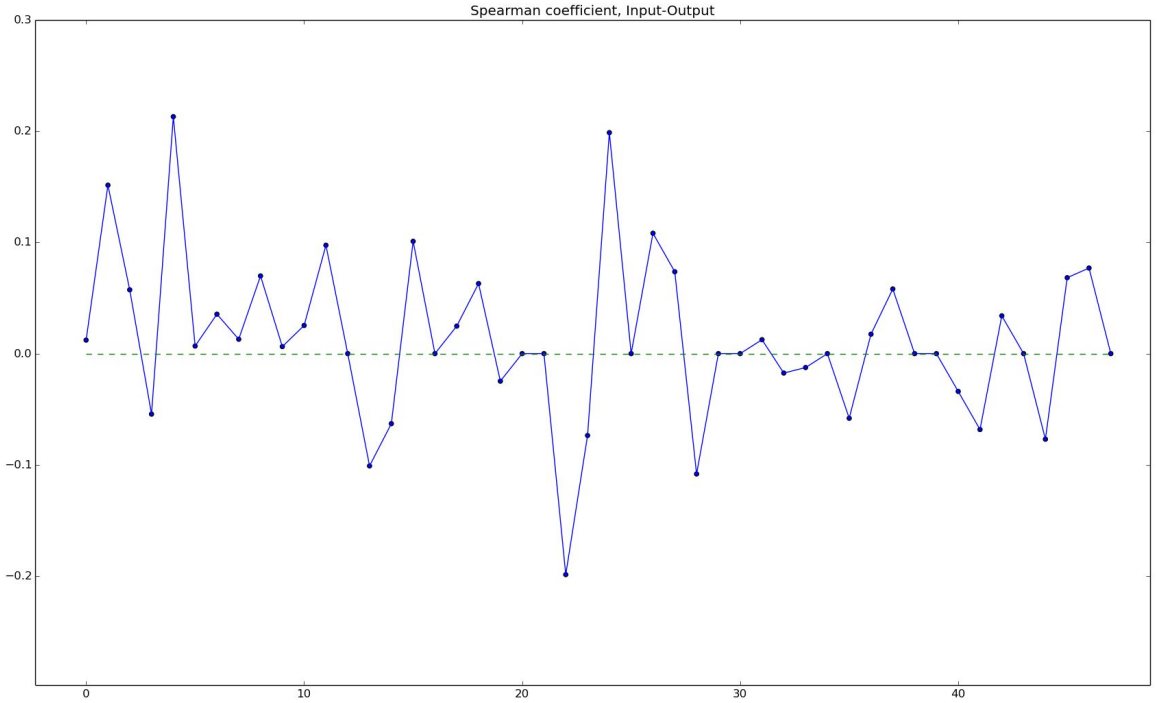
Correlation test

- Each of the peak and/or trough of the feature interacts with atmosphere at different heights
- But the features are so closely formed that it is expected have a common response layer
- Therefore, it will affect the temperature (T_avg) at an average height
- If the features are affecting the “T_avg”, then they should be well correlated to each other



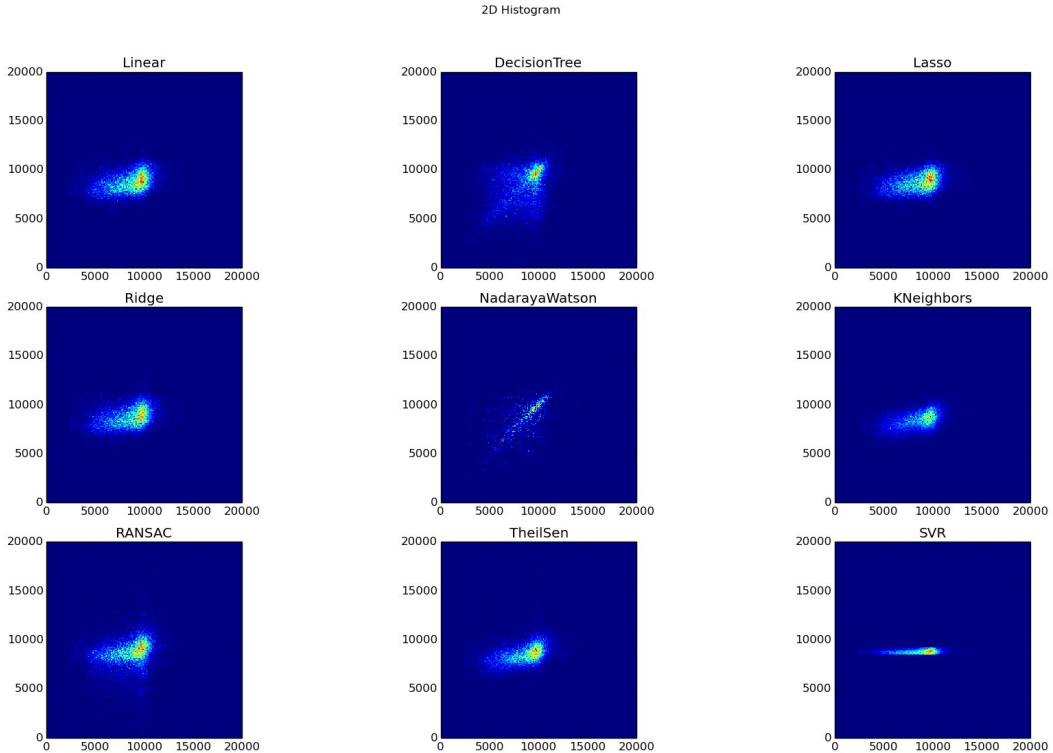
Correlation test II

- The “T_avg” is also correlated to the features that are used to predict it
- The plot shows that the output is correlated/anti-correlated to most of the features



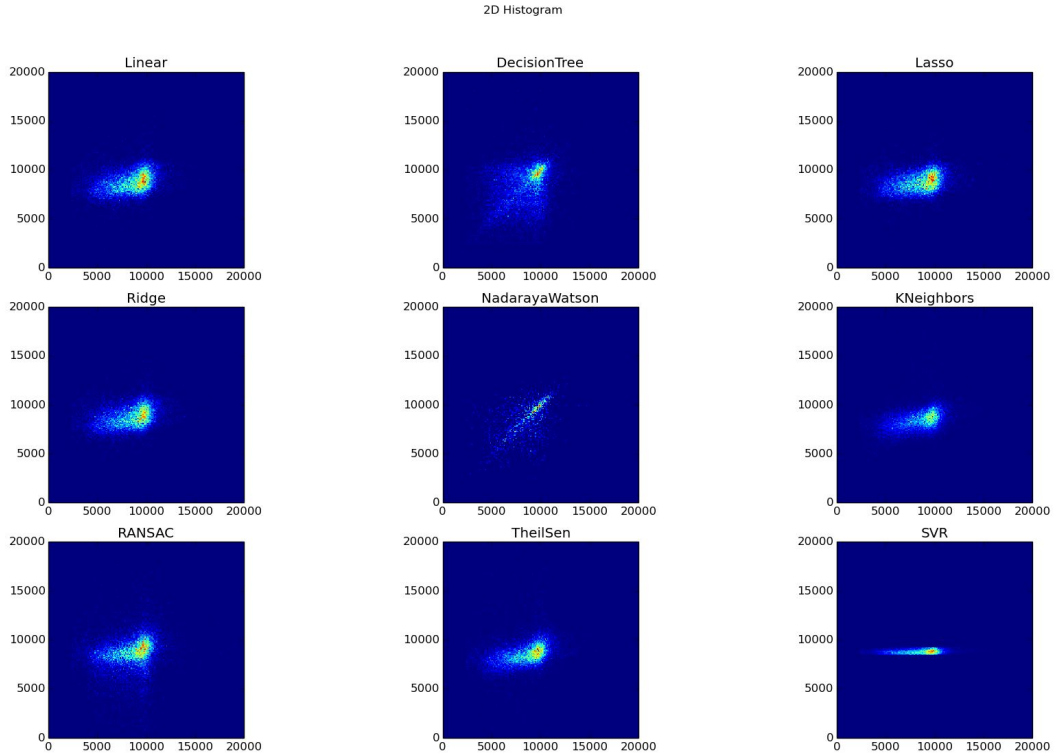
Regression

- **Steps:**
  - Clean the NAN from the data
  - Split the data points into training, cross-validation and test set
  - Apply each of the regression methods
  - Plot the 2D histogram of the true and predicted label



Regression II

- **2D Histogram:**
  - The plot shows how well each of the methods predicted the true values
  - The plot shows that most of the methods couldn't do so well
  - "Decision tree regression" and "Nadaraya-Watson regression" seems to give better results than others



Results

- The better methods are selected to go for a more difficult test, i.e., predicting a patch of the T_avg image
- The small box is predicted using the rest part of the image to train the method

Results II

- **Linear Regression:**
  - It is able to just predict the average value of the small box
- **KNeighbors (k=15):**
  - It gives better prediction than linear regression
  - The image is very grainy
  - However, it predicts the large scale features quite well

Results III

- **Nadaraya-Watson:**
  - The method couldn’t predict outputs for many points
  - But the values that it did predict are very close to true values
- **Decision Tree Regression:**
  - This method gave the best result
  - Along with the large scale features, it could predict also the small scale features of the image

Conclusion

- The regression methods can be used to predict the temperature at the average height
- A rough estimate of the atmosphere of the chromosphere can be determined
- Some of the methods can also predicts the large scale events
- For more accurate results, the MHD simulations become necessary. However, the predictions from the machine learning techniques can be used as initial guess values

Thank you!