[*Ruano et al.*, 2014] talks about building MLP and SVM method to classify seismic signals. They use sliding window to extract features from one station, and train a MLP and SVM method to classify different signals. For the SVM, they use an active learning method to retrain on the mis-classified results, and achieved better results. Besides, they also train a model with shorter time window which can be applied to EEW. Overall, it is a paper overlays the correct procedure to do signal classification. But a few things not correctly or clearly stated, for example, for the imbalanced dataset, how they deal with it.

[*Shirzaei et al.*, 2016] reports the surface uplift due to the injection in eastern Texas. Using time evolution InSAR images, they get a uplift of 3 mm/year over >8km area from the injection wells. They also build a poroelastic model to explain the uplift. They invested two wells, one in the west, and one in the east, and only the east shows the uplift. The west wells do not show the uplift, but accompanied by a sequence of earthquakes. They attribute this to the low compressibility of the rocks at the west wells. Therefore, the seismicity and the deformation behavior depends both on the injection activity and the local hydrogeological properties. Some interesting things from this paper: (1) Seismic activity increased even while the injection rates declined, owing to diffusion of pore pressure from earlier periods with higher injection rates. (2) Induced seismicity potential is suppressed where tight confining formations prevent pore pressure from propagating into crystalline basement rocks. (3) Over time, the increased pore pressure due to injection can spread to distances of many kilometers. (4) A localized increase in pore pressure shifts the circle (Mohr circle) to the left and changes its radius because of poroelastic strain, whereas a spatially uniform pore pressure increase only shifts the circle to the left until it touches the failure envelope.

[*Levander et al.*, 2011] proposed the hypothesis for the uplift of the Colorado plateau – delamination style convective lithospheric downwelling. The main method they used is from body wave tomography and receiver function. Based on the results from these two methods, they propose the process like this: The Colorado plateau lithosphere has been dydrated, and the small increase in density from the freezing metls, and the viscosicty reduction from hydration and advected heat, destabilizes the lithosphere and initiates a localized downwelling. The re-fertilized Colorado plateau mantle has been removed, delaminating the lowermost crust with it. The asthenosphere is invading the region from the beneath the drip and around the peripheries of the drip. They inferred that the lowermost crust involved in the dwonwelling has been modified by intrusion of basaltic melts that froze to produce high density eclogites. A series of these events have been removing the lithosphere from the Colorado plateau peripheries since the Farallon slab was removed 20-30 Myr ago, and causing the uplift.

[*Taira et al.*, 2009] reports the changes of fault-strength on the San Andreas fault at Parkfield by remote triggering. They argue that the seismicity of the repeating earthquakes at Parkfield have revealed a means of monitoring fault strength. For the Landers and Sumatra earthquake, the dynamic strain causes the changes in two manifestations: temporal variations in the properties of seismic scatters – probably reflecting the stress-induced migration of fluids – and systematic temporal variations in the characteristics of repeating-earthquake sequences that are most consistent with changes in fault strength. They also found the 2004 M6 Parkfield earthquake causing the two remote triggering different maybe due to: it damaged the fault zone by creating new fractures, it relieved most of the stress stored in the fault zone, and the absence of accompanying slip for the 2004 Sumatra earthquake can be explained by the low driving stress.

[*Bayrakci et al.*, 2016] reports the fault-controlled hydration of the upper mantle during continental rifting. They shoed the serpentinization at the rifted continental margin offshore from western Spain was probably initiated when the whole crust cooled to become brittle and deformation was focused along large normal faults. They use seismic tomography to image the 3D distribution of serpentinization in the mantle and find that the local volume of serpentinite beneath thinned, brittle crust is related to the amount of displacement along each fault. This implies that sea water reaches the mantle only when the faults are active. They also estimate the fluid flux along the faults and find it is comparable to that inferred for mid-ocean ridge hydrothermal systems. They conclude that brittle processes in the crust may ultimately control the global flux of sea water into the Earth. Some useful backgrounds can be found [here](https://en.wikipedia.org/wiki/Non-volcanic_passive_margins).

[*Lenardic*, 2017] gives a very nice review about a paper published in nature geoscience by Van Avendonk et al. The hypothesis proposed by the paper is that the changes of the earth’s internal energy cooling rate leaves a trace on the sea floor thickness. Oceanic crust forms dominantly by decompression melting of mantle rocks below mid-ocean ridges, so raised mantle temperatures can result in thicker crust forming at the ridge. Therefore, by examining the sea floor thickness, they can estimate the variations in mantle temperature through time. They used seismic data gathered over the past 40 years to determine the thickness of oceanic crust across the globe and show that, on average, the oceanic crust has thinned. They also notice that the cooling rate below the Atlantic and Indian mid-ocean ridges is about three times higher than that beneath the Pacific. Immediately prior to this time, the Atlantic and Indian oceanic basins were sites above which the supercontinent Pangaea resided. Therefore, they argue that the supercontinent has an insulating effect which the high temperature beneath it may cause the instability and initiates the break up of the supercontinent. Thus we see a transit signal in these places in the last 100 M years. The implications of such fluctuations go beyond internal Earth dynamics, it also link to the greenhouse gas released, since it has been argued that the greenhouse world that our planet experienced in the Cretaceous may be connected to a volcanic-tectonic forcing event associated with Pangaea’s break-up.

[*Mishra and Gordon*, 2016] talked about the debates between the rigid-plate and the shrinking-plate hypotheses, and try to provide more evidences to support the shrinking-plate hypothesis. They use model to predict the azimuth of the transform-fault distributed between 15 plate pairs, and test whether a significantly better fit to the data is obtained after correction for the predicted bias. The three key points they got from this paper are: (1) The shrinking-plate hypothesis predicts subtle differences in azimuths of right-lateral versus left-lateral transform faults; (2) Transform-fault azimuths observed globally indicate a statistically significant difference between right-lateral and left-lateral faults; (3) Transform faults do not precisely parallel plate motion, thus validating inferred quantifiable plate nonrigidity.

[*Zheng et al.*, 2014] reports a modified method to estimate absolute plate velocities from seismic anisotropy by correcting the correlated errors. There are two common methods to estimate the absolute velocity: trends of hot spot tracks, and seismic anisotropy from shear wave splitting. But the current seismic anisotropy method assumes the errors in the azimuths inferred from shear wave splitting are uncorrelated, which is they show find that residuals to azimuths from any one plate are strongly correlated with the residuals from the same plate. Therefore this paper proposed a new method for analyzing the azimuths inferred from shear wave splitting data. The key points of this paper are: Absolute plate velocities differ insignificantly from zero for eight plates and seismic anisotropy indicates plate motion direction if speed exceeds ~5 mm/a.

[*Niu et al.*, 2003] reports the evidence of crustal structural changes at Parkfield that detected from seismic data. They find a systematic temporal variation in the seismograms of repeat microearthquakes that occurred on the Parkfield segment of the San Andreas fault over the decade 1987 – 97. Their analysis reveals a change of the order of 10 m in the location of scatters which plausibly lie within the fault zone at a depth of ~3 km. The motion of the scatters is coincident, in space and time, with the onset of a well documented aseismic transient (deformation event). They speculate that this structural change is the result of a stress-induced redistribution of fluids in fluid-filled fractures caused by the transient event.

[*LeCun et al.*, 2015] gives a very nice review about Deep learning in nature. They first talked about the benefits of deep learning, the key advantage is that deep learning can learn good features automatically using a general-purpose learning procedure, thus avoid of feature engineering. Some key points from their reviews are:

1. Representation learning is a set of methods that allows a machine to be fed with raw data and to automatically discover the representations needed for detection or classification.
2. Stochastic gradient descent called stochastic because each small set of examples gives a noisy estimate of the average gradient over all examples.
3. At present, the most popular non-linear function is the rectified linear unit, which typically learns much faster in networks with many layers compare with tanh and sigmoid.
4. The hidden layers can be seen as distorting the input in a non-linear way so that categories become linearly separable by the last layer.
5. In practice, poor local minima are rarely a problem with large networks.
6. The unsupervised ‘pre-training’ can reduce the overfitting problem when use small datasets, and it is needed when you have a small dataset
7. The first major application of this pre-training approach was in speech recognition, and it was made possible by the advent of fast graphics processing units (GPU).
8. By 2012, versions of deep net were already being deployed in Android phones.
9. The convolutional neural network (ConvNets) was much easier to train and generalized much better than networks with full connectivity between adjacent layers.
10. ConvNets are designed to process data that come in the form of multiple arrays. There are 4 key ideas that take advantage of the properties of natural signals: local connections, shared weights, pooling and the use of many layers.
11. New regularization method is to use the dropout.
12. Recurrent neural networks used for sequential data by unfolding the time into a very deep nets.
13. RNN is not good to store information for very long time, therefore, the proposal of long short-term memory networks is to solve this problem.

[*Long et al.*, 2014] gives a review of the USArray and the introduction of the papers in the special issue from this array. It summaries the exciting results from USArray mainly from the study of the structure. Therefore it provides a quick way to know what the papers in this issue about. But I am thinking the other type of study for USArray, maybe I should spend more time on the Machine Learning and see if I can identify some interesting projects.

[*Goodfellow et al.*, 2014] proposed the generative adversarial nets, a very interesting idea. It has two models that compete with each other, one generative model and one discriminative model, that being trained together. The generative model (or generator) is trying to produce fake data to fool the discriminative model, while the discriminative model is try to distinguish whether the input data is from the real world data or the generator. Competition in this game drives both teams to improve their methods until the generator generates data points so real that the discriminative model can not distinguish it. The argument is even we don’t know what exactly the generative model works inside, if it can generate realistic sample, we can think it has a very nice internal mechanism to model the world. You can see [*Focus*, 2017] for an example to generate realistic samples in the astronomy study. Personally, I think this is really a cool method that I can use to generate realistic data without build a parametric model!

Bayrakci, G. et al. (2016), Fault-controlled hydration of the upper mantle during continental rifting, *Nat. Geosci.*, (March), 1–6, doi:10.1038/ngeo2671.

Focus, N. I. N. (2017), Astronomers explore uses for AI-generated images Hydrogen yet to prove it ’ s metal, *Nature*, 6–7.

Goodfellow, I. J., J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio (2014), Generative Adversarial Networks, , 1–9.

LeCun, Y., Y. Bengio, and G. Hinton (2015), Deep learning, *Nature*, *521*(7553), 436–444, doi:10.1038/nature14539.

Lenardic, A. (2017), PLATE TECTONICS A supercontinental boost, *Nat. Publ. Gr.*, *10*(1), 4–5, doi:10.1038/ngeo2862.

Levander, A., B. Schmandt, M. S. Miller, K. Liu, K. E. Karlstrom, R. S. Crow, C. T. A. Lee, and E. D. Humphreys (2011), Continuing Colorado plateau uplift by delamination-style convective lithospheric downwelling, *Nature*, *472*(7344), 461–465, doi:10.1038/nature10001.

Long, M. D., A. Levander, and P. M. Shearer (2014), An introduction to the special issue of Earth and Planetary Science Letters on USArray science, *Earth Planet. Sci. Lett.*, *1*(September 2014), 1–5, doi:10.1016/j.epsl.2014.06.016.

Mishra, J. K., and R. G. Gordon (2016), The rigid-plate and shrinking-plate hypotheses: Implications for the azimuths of transform faults, *Tectonics*, *35*(8), 1827–1842, doi:10.1002/2015TC003968.

Niu, F., P. G. Silver, R. M. Nadeau, and T. V. McEvilly (2003), Migration of seismic scatterers associated with the 1993 Parkfield aseismic transient event, *Nature*, *426*(6966), 544–548, doi:10.1038/nature02151.

Ruano, A. E., G. Madureira, O. Barros, H. R. Khosravani, M. G. Ruano, and P. M. Ferreira (2014), Seismic detection using support vector machines, *Neurocomputing*, *135*(January), 273–283, doi:10.1016/j.neucom.2013.12.020.

Shirzaei, M., W. L. Ellsworth, K. F. Tiampo, P. J. Gonzalez, and M. Manga (2016), Surface uplift and time-dependent seismic hazard due to fluid injection in eastern Texas, *Science (80-. ).*, *353*(6306), 1416–1419, doi:10.1126/science.aag0262.

Taira, T., P. G. Silver, F. Niu, and R. M. Nadeau (2009), Remote triggering of fault-strength changes on the San Andreas fault at Parkfield, *Nature*, *461*(7264), 636–639, doi:10.1038/nature08395.

Zheng, L., R. G. Gordon, and C. Kreemer (2014), Absolute plate velocities from seismic anisotropy: Importance of correlated errors, *J. Geophys. Res. Solid Earth*, *119*(9), 7336–7352, doi:10.1002/2013JB010902.