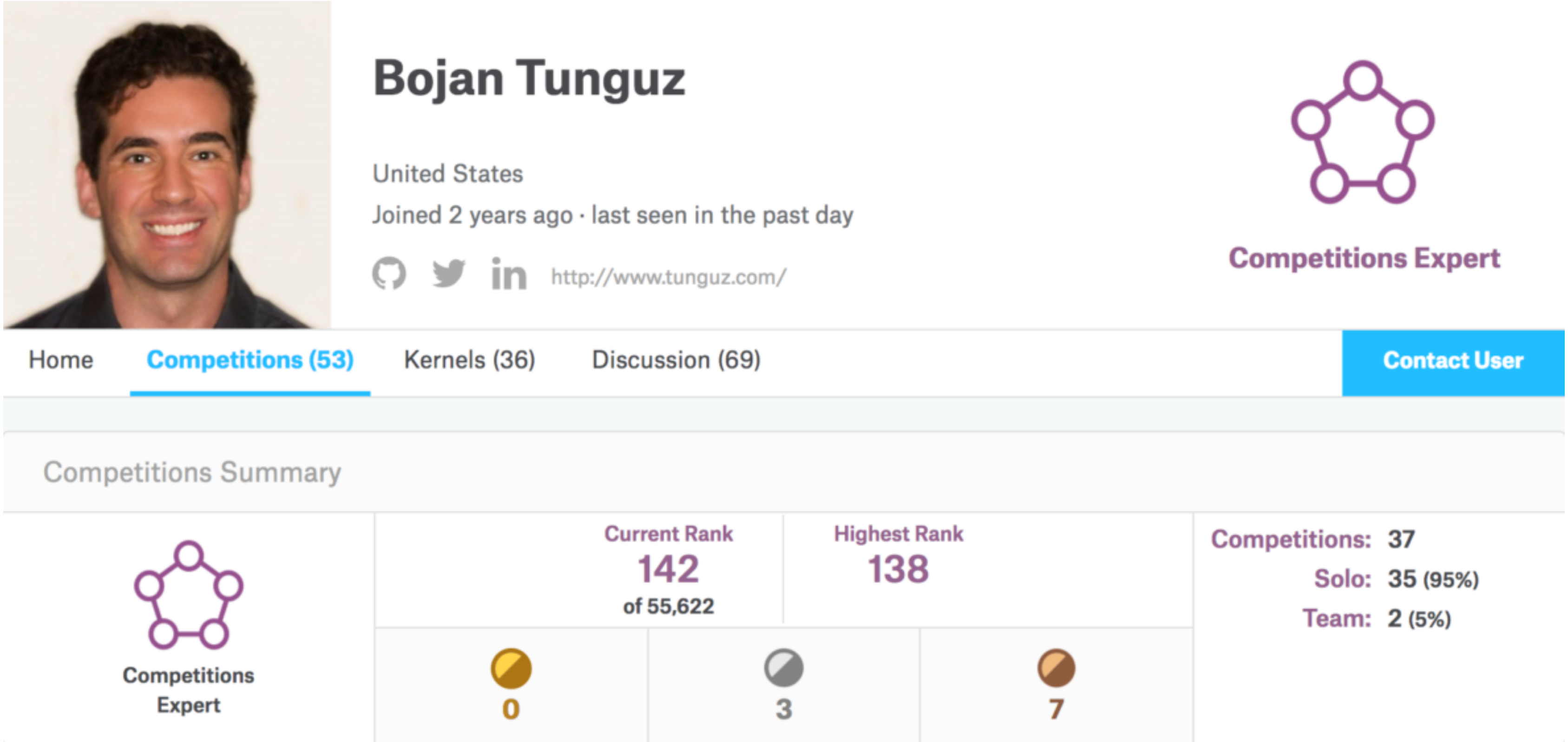
Dogs vs. Cats Redux Playground Competition, Winner's Interview: Bojan Tunguz

The Dogs versus Cats Redux: Kernels Edition playground competition revived one of our favorite "for fun" image classification challenges from 2013, Dogs versus Cats. This time Kaggle brought Kernels, the best way to share and learn from code, to the table while competitors tackled the problem with a refreshed arsenal including TensorFlow and a few years of deep learning advancements. In this winner's interview, Kaggler Bojan Tunguz shares his 4th place approach based on deep convolutional neural networks and model blending.

The basics

What was your background prior to entering this challenge?

I am a Theoretical Physicist by training, and have worked in Academia for many years. A few years ago I came across some really cool online machine learning courses, and fell in love with that field. I've been doing some freelancing data science and machine learning work for a while, and now I work for a FinTech startup.



Bojan Tunguz

United States

Joined 2 years ago · last seen in the past day

http://www.tunguz.com/

Competitions Expert

Competitions Summary

Competitions Expert

Current Rank 142 of 55,622

Highest Rank 138

0 3 7

Competitions: 37

Solo: 35 (95%)

Team: 2 (5%)

Bojan Tunguz on Kaggle.

Do you have any prior experience or domain knowledge that helped you succeed in this competition?

When I was growing up my family owned several cats. I also watch a lot of online cat videos, so I feel I have a pretty good idea of what cats look like and how they differ from dogs.

So far I have not had any “official” experience in computer vision outside of Kaggle competitions. However, I have competed pretty successfully in a few other image recognition/categorization competitions, and I count this as one of my core machine learning competencies.

How did you get started competing on Kaggle?

I’ve been hearing about Kaggle for years, but finally decided to take the plunge and start competing about a year and a half ago (September 2015). I was initially apprehensive about competing on such a high level, but Kaggle’s community, kernels, discussions, etc., were very useful and helpful in getting me up to speed.

What made you decide to enter this competition?

There are several things I liked about the Dogs vs. Cats Redux competition that made me want to spend a lot of my time on it. As I already mentioned, I like image categorization competitions, and on average I do pretty well on them. This competition also seemed as “pure” of a machine learning categorization problems as they come: just two, perfectly balanced categories, with enough data to build sophisticated models. As I tried a few early solutions, the problem seemed pretty “blendable,” i.e. blending solutions from different models would generally improve the public leaderboard score. This suggested to me that building advanced “meta” models would be relatively straightforward to do. The competition also started at the time when I didn’t see many other interesting competitions on Kaggle. I also liked the fact that this was a repeat of a competition that was hosted on Kaggle before, so it was interesting to compare the methods and solutions from that competition, and see how far the field of image classification has progressed in just a few years. Finally, since this was a “Playground” competition that ran for about half a year, it gave me ample time and opportunity to try out different strategies and refine my image classification skills without the added pressure of being one of the “Featured” competitions.

Let’s get technical

Did any past research or previous competitions inform your approach?

Image classification problems have by now become almost commoditized, and there are a lot of good papers, tools, and software libraries the help you get started. The Deep Learning community has been generously offering many of their pretrained models for free, and these would be prohibitively expensive and time consuming to train “from scratch”. I have also benefitted from my experience with other previous and current image classification competitions (Yelp Restaurant Photo Classification, State

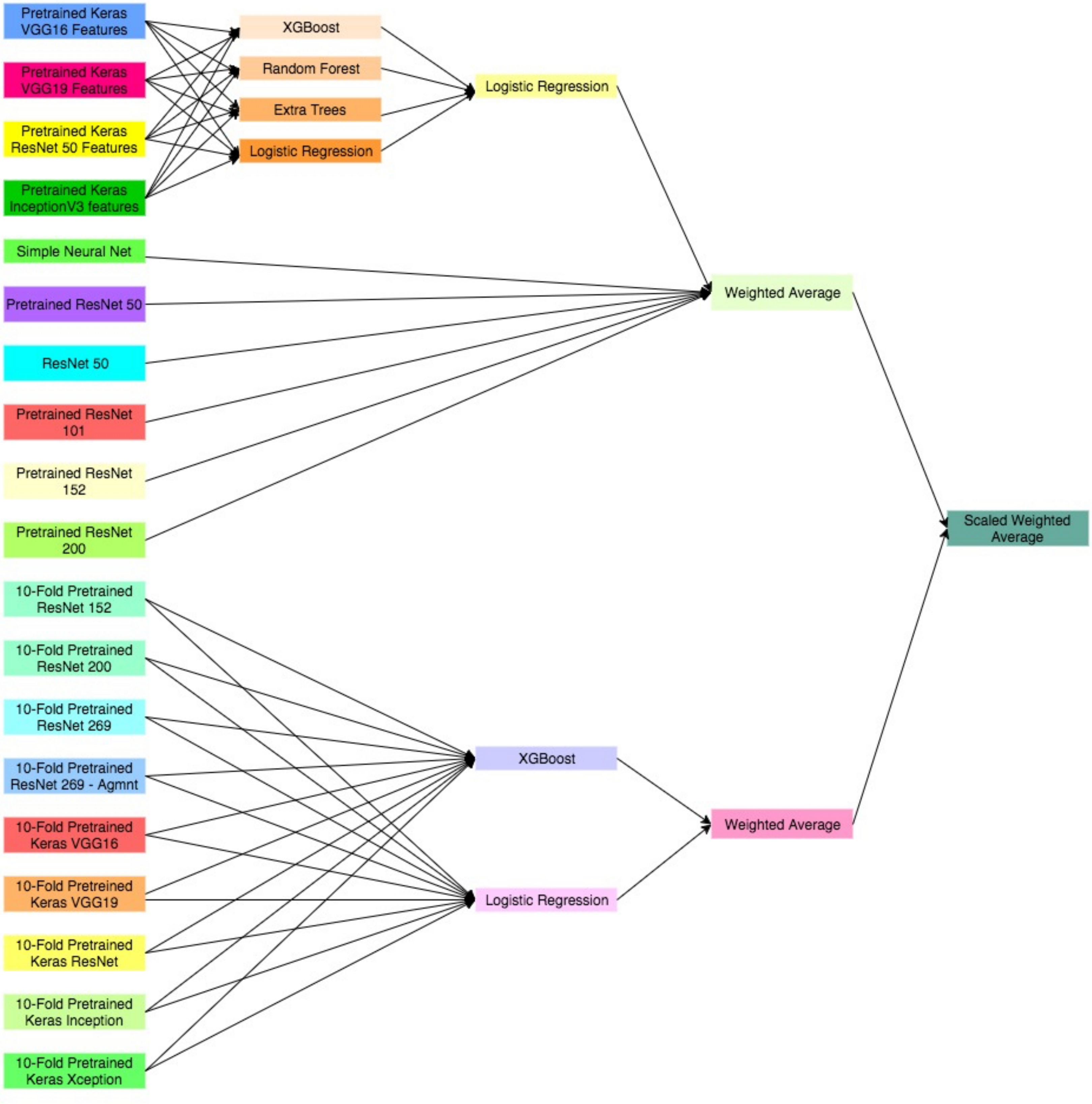
Farm Distracted Driver Detection, Nature Conservancy Fisheries Monitoring, etc.) which have greatly helped with refining my workflow.

What preprocessing and feature engineering did you do?

I spent relatively little time on preprocessing and feature engineering. I had split data for various cross validation folds on disk, in order to ensure the full consistency across multiple models/machines, as well as for easier access by various command line tools that I used. For one of my models I've done a lot of image augmentation - cropping, shearing, rotating, flipping, etc.

What supervised learning methods did you use?

Just like with most other image recognition/classification problems, I have completely relied on Deep Convolutional Neural Networks (DCNN). I have built a simple convolutional neural network (CNN) in Keras from scratch, but for the most part I've relied on out-of-the-box models: VGG16, VGG19, Inception V3, Xception, and various flavors of ResNets. My simple CNN managed to get the score in the 0.2x range on the public leaderboard (PL). My best models that I build using features extracted by applying retrained DCNNs got me into the 0.06x range on PL. Stacking of those models got me in the 0.05x range on PL. My single best fine-tuned DCNN got me to 0.042 on PL, and my final ensemble gave me the 0.35 score on PL. My ensembling diagram can be seen below:


Pretrained Keras VGG16 Features
Pretrained Keras VGG19 Features
Pretrained Keras ResNet 50 Features
Pretrained Keras InceptionV3 features
XGBoost
Random Forest
Extra Trees
Logistic Regression
Logistic Regression
Simple Neural Net
Pretrained ResNet 50
ResNet 50
Pretrained ResNet 101
Pretrained ResNet 152
Pretrained ResNet 200
Weighted Average
Scaled Weighted Average
10-Fold Pretrained ResNet 152
10-Fold Pretrained ResNet 200
10-Fold Pretrained ResNet 269
10-Fold Pretrained ResNet 269 - Agmnt
10-Fold Pretrained Keras VGG16
10-Fold Pretreined Keras VGG19
10-Fold Pretrained Keras ResNet
10-Fold Pretrained Keras Inception
10-Fold Pretrained Keras Xception
XGBoost
Logistic Regression
Weighted Average

Which tools did you use?

I have primarily used Keras and a Facebook implementation of pretrained ResNets. The latter is written in Torch, so as I am not proficient in Lua, I had to develop all sorts of hacks to get the output of command line tools into my main Python scripts. I have also used OpenCV, XGBoost and sklearn for image manipulation and stacking.

How did you spend your time on this competition?

I have not done much feature engineering for this competition. Only one of my cross-validation models used significantly augmented images for its training input. I would say that for this competition I spent about 5% of my time on feature engineering, and the rest on machine learning.

What does your hardware setup look like?

I've built my own Ubuntu desktop box specifically for machine learning projects - i7 Intel processor, ASUS motherboard, 32 GB of RAM and dual NVIDIA GTX 970/960 cards. I've built and trained most of my models for this competition on that machine. Recently I've been able to avail of a System76 laptop with 64 GB of RAM and NVIDIA GTX 1070 GPU, but I have not used it for any of my most advanced models.

What was the run time for both training and prediction of your winning solution?

The most elaborate model that I used for this competition was a 10-fold CV 269-layer deep ResNet. It took about 15 hours to train each fold on my machine, so that translates into about 6 days training. The prediction phase was about 20 minutes per fold, so about three and a half hours total. As I mentioned above, I've viewed this competition as a good practice for learning how to fine tune neural networks for image recognition/classification problems, and and over the course of its duration I've spent many weeks worth of computational time on various different models.

Words of wisdom

What have you taken away from this competition?

Given enough clean, well-defined image data, deeper the CNN model the better.

Looking back, what would you do differently now?

I would try harder and start earlier to look into training really deep neural networks from scratch. I've been able to train a ResNet 50 from scratch, and it outperformed the fine tuned pertained model. However, I have not been able to do the same with the deeper NNs - my training was stuck in a rut. I would also invest more time in the localization of cats and dogs in images, and maybe even train a separate NN for that task. I would also look into getting additional data form other sources, since this seems to be allowed by the competition rules.

Do you have any advice for those just getting started in data science?

Just do it! If you are interested in data science, just start reading available resources, taking online classes, and, of course check out Kaggle competitions and tutorials. Regardless of your previous level of competence in coding and statistics, I believe the best way to get started with data science is just to take the plunge and start working on some projects. Learn R and/or Python, the two most popular languages with data scientists. Look at other people's code, and then play with it and modify it to see what happens. Don’t be intimidated by the complex-sounding terms and algorithms.

I have taken several different online courses, and I would recommend the ones offered through Coursera and Udacity. Check out the Kaggle tutorial competitions: Digit Recognizer, Titanic and House Prices. They provide a lot of useful kernels that you can play with and modify. Go through Kaggle discussion boards - they too have tons of useful information. Don’t hesitate to ask questions - we’ve all been “noobs” at some point, and it was thanks in no small part to those who were patient enough to explain some “simple” concepts to us that we finally got where we are now.

Bio

Bojan Tunguz works for ZestFinance as a Machine Learning Modeler. He has been involved in data science and machine learning for about 3 years. He holds BS and MS degrees in Physics and Applied Physics from Stanford University, and a Ph.D. in Physics from University of Illinois at Urbana-Champaign. He currently doesn’t own any dogs or cats, but hopes that this state of affairs will not long endure.