

Want better deep learning models? Take the [FREE Mini-Course](#)

Search...



A Gentle Introduction to Transfer Learning for Deep Learning

by Jason Brownlee on December 20, 2017 in [Better Deep Learning](#)

Tweet

[Share](#)

[Share](#)

[G+](#)

Transfer learning is a machine learning method where a model developed for a task is reused as the starting point for a model on a second task.

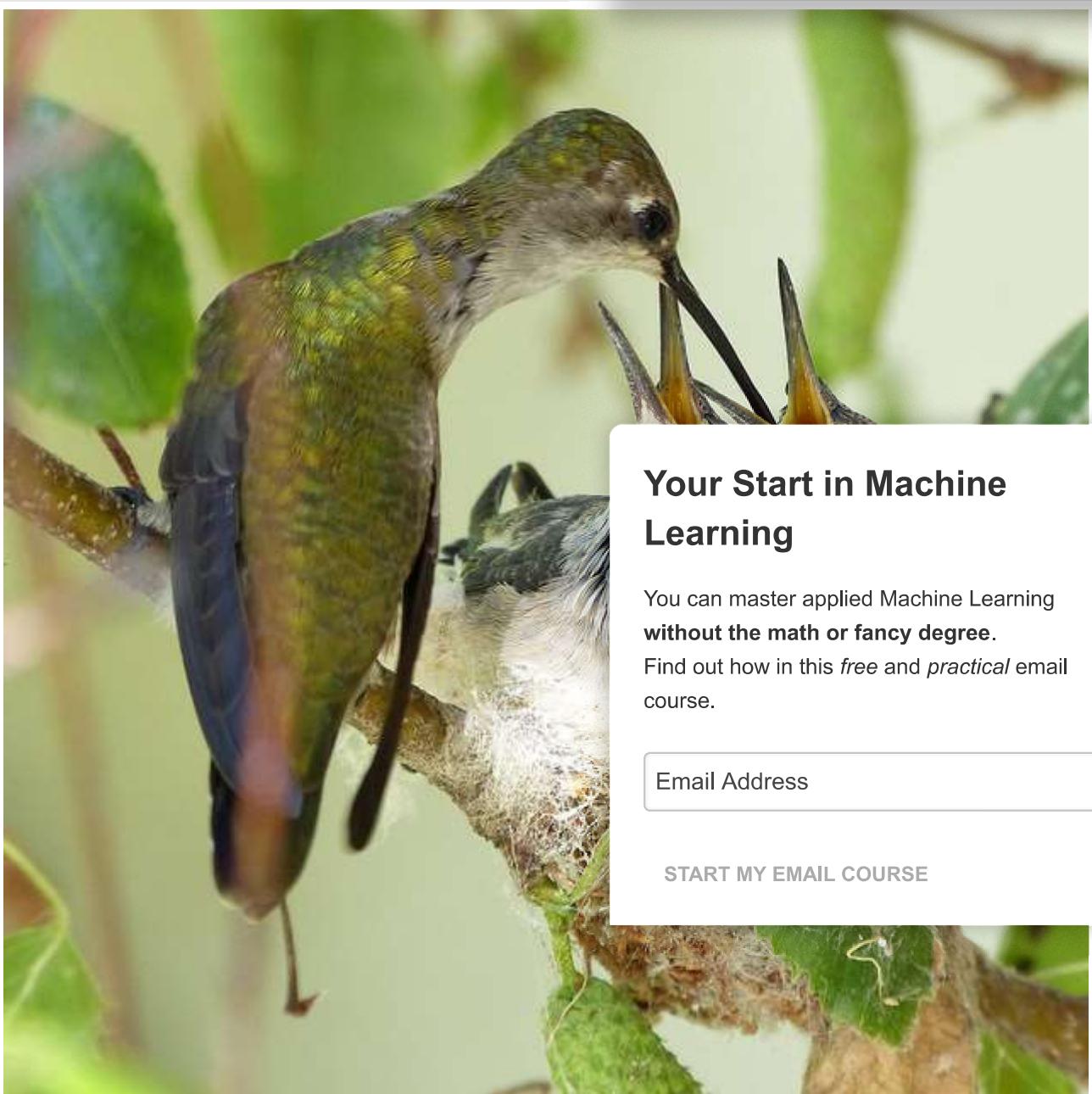
It is a popular approach in deep learning where pre-trained models are used as the starting point on computer vision and natural language processing tasks given the vast compute and time resources required to develop neural network models on these problems and from the huge jumps in skill that they provide on related problems.

In this post, you will discover how you can use transfer learning to speed up training and improve the performance of your deep learning model.

After reading this post, you will know:

- What transfer learning is and how to use it.
- Common examples of transfer learning in deep learning.
- When to use transfer learning on your own predictive modeling problems.

Let's get started.



Your Start in Machine Learning

You can master applied Machine Learning **without the math or fancy degree.**

Find out how in this *free* and *practical* email course.

[START MY EMAIL COURSE](#)

A Gentle Introduction to Transfer Learning with Deep Learning
Photo by Mike's Birds, some rights reserved.

What is Transfer Learning?

Transfer learning is a machine learning technique where a model trained on one task is re-purposed on a second related task.

“ Transfer learning and domain adaptation refer to the situation where what has been learned in one setting ... is exploited to improve generalization in another setting

— Page 526, [Deep Learning](#), 2016.

Transfer learning is an optimization that allows rapid progress or improved performance when modeling the second task.

“ Transfer learning is the improvement of learning in a new related task that has already been learned.

Your Start in Machine Learning

Transfer learning is related to problems such as multi-task learning and concept drift and is not exclusively an area of study for deep learning.

Nevertheless, transfer learning is popular in deep learning given the enormous resources required to train deep learning models or the large and challenging datasets on which deep learning models are trained.

Transfer learning only works in deep learning if the model features learned from the first task are general.

In transfer learning, we first train a base network on a base dataset and tasks, and then we re-purpose the learned features, or transfer them, to a second target network on a different dataset and tasks. This process will tend to work if the features are general, i.e., if they are shared across tasks instead of specific to the base task.

— How transferable are features in deep neural networks?

This form of transfer learning used in deep learning is called inductive transfer. The bias of the learned models (model bias) is narrowed in a beneficial way by using a model learned on a related task.

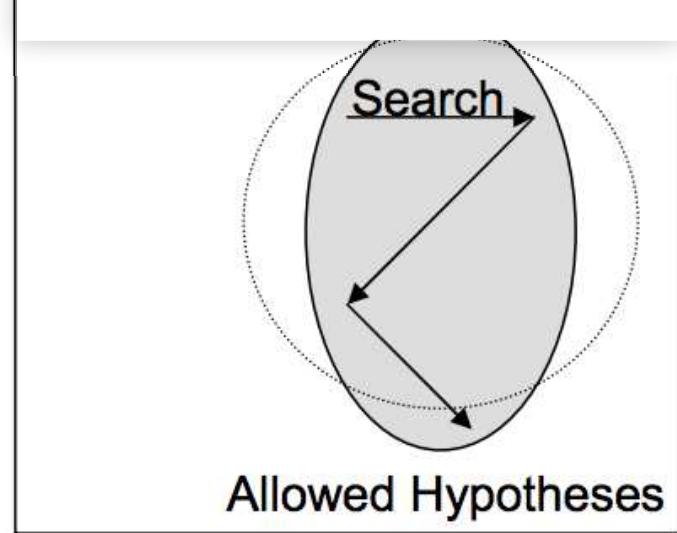
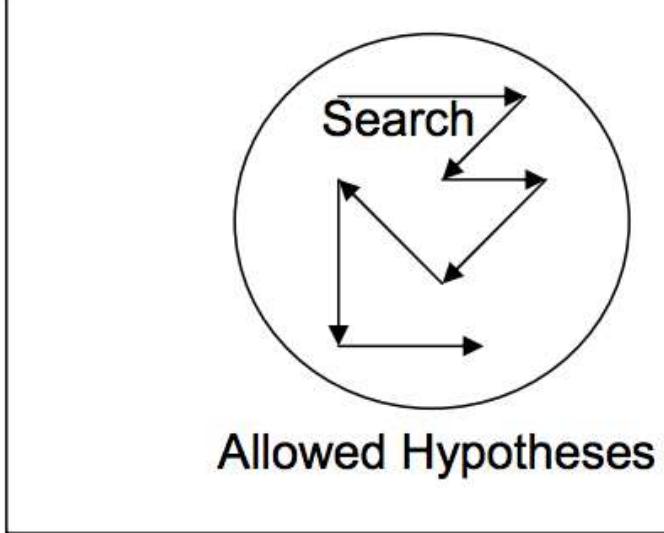
Your Start in Machine Learning

You can master applied Machine Learning without the math or fancy degree.

Find out how in this free and practical email course.

Email Address

START MY EMAIL COURSE



Depiction of Inductive Transfer
Taken from “Transfer Learning”

Want Better Results with Deep Learning?

Take my free 7-day email crash course now (with sample code).

Click to sign-up and also get a free PDF

Your Start in Machine Learning

How to Use Transfer Learning?

You can use transfer learning on your own predictive modeling problems.

Two common approaches are as follows:

1. Develop Model Approach
2. Pre-trained Model Approach

Develop Model Approach

1. **Select Source Task.** You must select a related predictive modeling task where there is some relationship in the input data, output data, and/or the output data.
2. **Develop Source Model.** Next, you must develop a skillful model, perhaps even a naive model to ensure that some feature learning has been achieved.
3. **Reuse Model.** The model fit on the source task can then be reused on the second task of interest. This may involve using all or parts of the model.
4. **Tune Model.** Optionally, the model may need to be adapted or tuned on the task of interest.

Your Start in Machine Learning

You can master applied Machine Learning without the math or fancy degree.

Find out how in this *free* and *practical* email course.

[START MY EMAIL COURSE](#)

Pre-trained Model Approach

1. **Select Source Model.** A pre-trained source model is chosen from available models. Many research institutions release models on large and challenging datasets that may be included in the pool of candidate models from which to choose from.
2. **Reuse Model.** The model pre-trained model can then be used as the starting point for a model on the second task of interest. This may involve using all or parts of the model, depending on the modeling technique used.
3. **Tune Model.** Optionally, the model may need to be adapted or refined on the input-output pair data available for the task of interest.

This second type of transfer learning is common in the field of deep learning.

Examples of Transfer Learning with Deep Learning

Let's make this concrete with two common examples of transfer learning with deep learning models.

Transfer Learning with Image Data

It is common to perform transfer learning with predictive modeling problems that use image data as input.

This may be a prediction task that takes photographs or video data as input.

For these types of problems, it is common to use a deep learning model pre-trained for a large and challenging image classification task such as the ImageNet 1000-class photograph classification task.

The research organizations that develop models for this competition and do well often release their final model under a permissive license for reuse. These models can take days or weeks to train on modern hardware.

These models can be downloaded and incorporated directly into new models that expect image data as input.

Three examples of models of this type include:

- Oxford VGG Model
- Google Inception Model
- Microsoft ResNet Model

For more examples, see the [Caffe Model Zoo](#) where more pre-trained models are shared.

This approach is effective because the images were trained on a make predictions on a relatively large number of classes, in turn, features from photographs in order to perform well on the problem

In their Stanford course on Convolutional Neural Networks for Vis choose how much of the pre-trained model to use in your new mo

 [Convolutional Neural Networks] features are more generic in later layers

— Transfer Learning, CS231n Convolutional Neural Networks for

Transfer Learning with Language Data

It is common to perform transfer learning with natural language processing problems that use text as input or output.

For these types of problems, a word embedding is used that is a mapping of words to a high-dimensional continuous vector space where different words with a similar meaning have a similar vector representation.

Efficient algorithms exist to learn these distributed word representations and it is common for research organizations to release pre-trained models trained on very large corpora of text documents under a permissive license.

Two examples of models of this type include:

- Google's word2vec Model
- Stanford's GloVe Model

These distributed word representation models can be downloaded and incorporated into deep learning language models in either the interpretation of words as input or the generation of words as output from the model.

In his book on Deep Learning for Natural Language Processing, Yoav Goldberg cautions:

 ... one can download pre-trained word vectors that were trained on very large quantities of text [...] differences in training regimes and underlying corpora have a strong influence on the resulting representations, and that the available pre-trained representations may not be the best choice for [your] particular use case.

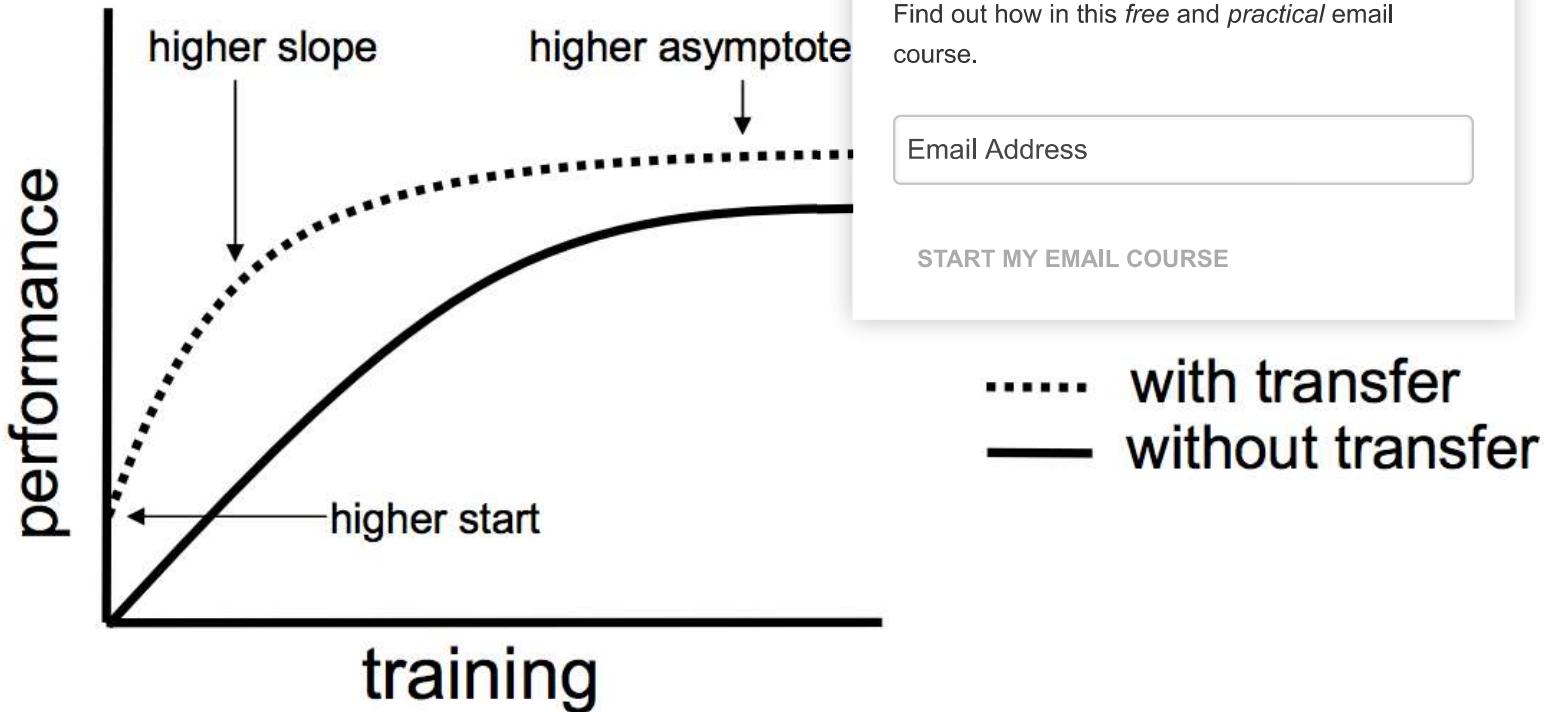
When to Use Transfer Learning?

Transfer learning is an optimization, a shortcut to saving time or getting better performance.

In general, it is not obvious that there will be a benefit to using transfer learning in the domain until after the model has been developed and evaluated.

Lisa Torrey and Jude Shavlik in their chapter on transfer learning describe three possible benefits to look for when using transfer learning:

1. **Higher start.** The initial skill (before refining the model) on the task.
2. **Higher slope.** The rate of improvement of skill during training would be.
3. **Higher asymptote.** The converged skill of the trained model.



Three ways in which transfer might improve learning.

Taken from "Transfer Learning".

Ideally, you would see all three benefits from a successful application of transfer learning.

It is an approach to try if you can identify a related task with abundant data and you have the resources to develop a model for that task and reuse it on your own problem, or there is a pre-trained model available that you can use as a starting point for your own model.

On some problems where you may not have very much data, transfer learning can enable you to develop skillful models that you simply could not develop in the absence of transfer learning.

The choice of source data or source model is an open problem and may require domain expertise and/or intuition developed via experience.

Further Reading

This section provides more resources on the topic if you are looking to go deeper.

Books

- Deep Learning, 2016.
- Neural Network Methods in Natural Language Processing, 2017.

Papers

- A survey on transfer learning, 2010.
- Chapter 11: Transfer Learning, Handbook of Research on Machine Learning, 2011.
- How transferable are features in deep neural networks?

Pre-trained Models

- Oxford VGG Model
- Google Inception Model
- Microsoft ResNet Model
- Google's word2vec Model
- Stanford's GloVe Model
- Caffe Model Zoo

Articles

- Transfer learning on Wikipedia
- Transfer Learning – Machine Learning’s Next Frontier, 2017.
- Transfer Learning, CS231n Convolutional Neural Networks for Visual Recognition
- How does transfer learning work? on Quora

Summary

In this post, you discovered how you can use transfer learning to speed up training and improve the performance of your deep learning model.

Specifically, you learned:

- What transfer learning is and how it is used in deep learning.
- When to use transfer learning.
- Examples of transfer learning used on computer vision and natural language processing tasks.

Do you have any questions?

Ask your questions in the comments below and I will do my best to answer.

Your Start in Machine Learning

You can master applied Machine Learning without the math or fancy degree. Find out how in this *free* and *practical* email course.

Email Address

START MY EMAIL COURSE

Develop Better Deep Learning Models Today!

Your Start in Machine Learning

Train Faster, Reduce Overfitting, and Ensembles

Better Deep Learning

Train Faster, Reduce Overfitting,
and Make Better Predictions

Jason Brownlee

MACHINE
LEARNING
MASTERY



Tweet

Share

Share

G+



About Jason Brownlee

Jason Brownlee, PhD is a machine learning specialist who teaches machine learning methods via hands-on tutorials.

[View all posts by Jason Brownlee →](#)

...with just a few lines of python code

Discover how in my new Ebook:

Better Deep Learning

It provides **self-study tutorials** on topics like: *weight decay, batch normalization, dropout, model stacking* and much more...

Bring better deep learning to your projects!

Skip the Academics. Just Results.

[Click to learn more](#)

Your Start in Machine Learning

You can master applied Machine Learning **without the math or fancy degree**.

Find out how in this *free and practical* email course.

Email Address

[START MY EMAIL COURSE](#)

◀ A Gentle Introduction to Exploding Gradients in Neural Networks

Why Applied Machine Learning Is Hard ▶

50 Responses to *A Gentle Introduction to Transfer Learning for Deep Learning*



HandsOnML December 20, 2017 at 2:42 pm #

REPLY ↗

As always, well written and insightful article. The additional resources you provide on this topic are also very helpful! Thanks.



Jason Brownlee December 20, 2017 at 3:50 pm #

REPLY ↗

Thanks.

Eric H December 22, 2017 at 5:32 am #

Your Start in Machine Learning



You repeated a sentence early in the article.

"Transfer learning only works in deep learning if the model features learned from the first task are general."



Jason Brownlee December 22, 2017 at 5:38 am #

REPLY ↩

Thanks, fixed!



Norma A. January 5, 2018 at 5:39 pm #

REPLY ↩

Hello Dr. Jason. Thank you so much for sharing clearly an ML and Deep Learning.

Reviewing the pre-trained model references, it appears that Google checked at <https://github.com/tensorflow/models/tree/master/inception>. Do you have idea if this is now placed at some other location? Than



Jason Brownlee January 6, 2018 at 5:52 am #

Your Start in Machine Learning



You can master applied Machine Learning without the math or fancy degree.

Find out how in this *free* and *practical* email course.

Email Address

START MY EMAIL COURSE



F. Ameen January 29, 2018 at 4:12 pm #

REPLY ↩

Hello Jason,

I am a big fan of your work. I wanted to ask that how to choose a pre-trained model for my specific problem. for example what would be a good pre-trained model that would work well on handwritten arabic alphabet classification?



Jason Brownlee January 30, 2018 at 9:47 am #

REPLY ↩

Perhaps a model trained on other character data?



F. Ameen January 31, 2018 at 2:33 am #

REPLY ↩

yes. because all keras pre-trained models require input size (244, 244, 3) while my data has (64, 64, 1) dimensions. also imagenet and character classification are very different problems. Most of the imagenet features are redundant for my problem.

Jason Brownlee January 31, 2018 at 9:47 am #

Your Start in Machine Learning



I believe the models could be adapted and may still add value.



Reemo A. February 22, 2018 at 7:06 am #

REPLY ↩

Many thanks for the good explanation..Can i ask you about your suggestion for the most suitable pre-trained model for image super resolution..how can i select it ?



Jason Brownlee February 22, 2018 at 11:22 am #

REPLY ↩

Sorry, I have not worked on super resolution. I don't have much time to do research. I can suggest you to do a google search?



Cat Chenal May 28, 2018 at 6:49 am #

Here is an implementation of image deblurring using Convolutional adversarial networks [https://github.com/erikbernadotte/adversarial-networks-image-deblurring-45e3ab6977b5]. Asking your question to its author might be fruitful. Good luck.

Your Start in Machine Learning

You can master applied Machine Learning **without the math or fancy degree**. Find out how in this *free* and *practical* email course.

Email Address

START MY EMAIL COURSE

X



Jason Brownlee May 28, 2018 at 2:32 pm #

REPLY ↩

Thanks for sharing.



Ahmad Raza April 17, 2018 at 1:32 pm #

REPLY ↩

Thanks! I have understand this concept. I want to ask a question that, for example there is a pre-trained (OCR) model for English language, and I want to transfer this model into (OCR) for my local language which has completely different alphabets than English. It would help me or not ?



Jason Brownlee April 17, 2018 at 2:55 pm #

REPLY ↩

I expect it would help.



Nastaran April 27, 2018 at 1:44 pm #

REPLY ↩

Just wondering if applying transfer learning is a solution to dealing with small data-set size?

Your Start in Machine Learning



Jason Brownlee April 27, 2018 at 2:28 pm #

REPLY ↩

It sure could be, nice one!



Lena April 27, 2018 at 7:51 pm #

REPLY ↩

Hi Jason, thank you for this article! I wonder if one could fine-tune a model like resnet, trained for classification, for the purpose of say image filtering? Like local contrast enhancement or denoise etc. What you think? Thanks



Jason Brownlee April 28, 2018 at 5:27 am #

Sounds like a good idea, try it. Let me know how you get on.



Lena April 28, 2018 at 2:45 pm #

Ok, thanks!



Vijay Ravi May 5, 2018 at 11:37 am #

Thanks! Helped a lot 😊



Jason Brownlee May 6, 2018 at 6:22 am #

REPLY ↩

I'm glad to hear that.



Ahmed May 25, 2018 at 10:14 am #

REPLY ↩

Many thanks

I want to ask if I have PDF content for example drawings and I need model to train I will give the model what is that geometry which pre trained model you can recommend?



Jason Brownlee May 25, 2018 at 2:52 pm #

REPLY ↩

Perhaps you might be able to use an existing computer vision model. It would be cheap to try and evaluate the results.

Your Start in Machine Learning



Nastaran June 21, 2018 at 11:29 am #

REPLY ↗

Just wondering what is the difference between transfer learning and unsupervised pre-training?



Jason Brownlee June 21, 2018 at 4:57 pm #

REPLY ↗

They are orthogonal ideas. In theory, could use transfer learning with a supervised or an unsupervised learning problem.



afan June 23, 2018 at 6:41 pm #

Hello sir i need your Help Please if you can .



Jason Brownlee June 24, 2018 at 7:31 am #

What do you need help with?



JG July 24, 2018 at 6:49 am #

I got an idea for transfer learning, I want to share for your opinion but, in the physical-mathematical domain. This is my analogy or equivalence (lets call it JGs approach). For example, what about if in order to solve a physical problem (in our case e.g. an image classification), I approach it with some mathematical description such as partial derivatives, Laplace operator, etc. (in our case this could be the neural model architecture, with convolutions, fully connected , data_augmentation, ...that we are intend to apply, etc). But because this math are very complexes with only know some few particular solutions under some initial o boundary values (and this could be the pre-trained weights in our ML or DL issue)...son under OTHER boundary or initial or complexes conditions we do not have any (of these analytical) solutions (this could be the case with other image classification or other datasets inputs) ...so we can use the pretty well known solution an try to apply for our problem case, where we intend to obtain similars solutions (this is now the transfer learning repurpose in order to search for similar or analogous solutions)...what do you think Mr. Brownlee about the analogy? Of course, as in real life, probably, from time to time, the solutions to our problems are radically different to the ones we try to reused (model or weight for radically different images sets or classification), ..but at least, the transfer learning essay serves as initial inspiration:-))



Jason Brownlee July 24, 2018 at 2:29 pm #

REPLY ↗

Not sure I follow, sorry.

Are you able to simplify the explanation?

Your Start in Machine Learning

You can master applied Machine Learning without the math or fancy degree.

Find out how in this *free* and *practical* email course.

Email Address

START MY EMAIL COURSE

Your Start in Machine Learning



PG August 7, 2018 at 9:39 pm #

REPLY ↗

Actually, Very vague explanation.. I read it twice. but still I don't know how to use it in Keras..
Do you have any short simple example with Keras?



Jason Brownlee August 8, 2018 at 6:19 am #

REPLY ↗

Sorry to hear that.

I show how to reuse the VGG model for photo captioning (but it is not short):

<https://machinelearningmastery.com/develop-a-deep-learning-project/>



C August 12, 2018 at 3:30 am #

Thanks for this. Lovely introduction to the area.

Do you know of any papers using the "Develop Model Approach"? I specifically.



Jason Brownlee August 12, 2018 at 6:35 am #

Your Start in Machine Learning

You can master applied Machine Learning **without the math or fancy degree**.

Find out how in this *free* and *practical* email course.

Email Address

START MY EMAIL COURSE



C August 12, 2018 at 7:00 pm #

REPLY ↗

I have done so but approaches seem to be either about the pre-trained model method or some method which leverages the knowledge that two domains are related and so adds an extra step to the process, such as dimensionality reduction.

The two survey papers also do not mention a simple Develop Model Approach.



Juma Ally September 1, 2018 at 11:34 am #

REPLY ↗

Thank you for introduction about transfer learning.

Is it possible to use this model for the calling records data with the attributes likes(source call, dist call, start time, duration and locations) for prediction such as user mobility prediction or characterize traffic on base station etc.?



Jason Brownlee September 2, 2018 at 5:26 am #

REPLY ↗

Your Start in Machine Learning

I would recommend this process when working on a new predictive modeling problem:
<https://machinelearningmastery.com/start-here/#process>



sana khan September 24, 2018 at 8:55 pm #

REPLY ↩

Hi, It's me Mr. Sana Ullah Khan, I want to work in detection of malignant cells in breast cytology images, I am confused that how I can use a CNN with transfer learning for detection and classification of malignant cells. I sketch one diagram, as well as i, want to share with you that how I can work. guide me, please



Jason Brownlee September 25, 2018 at 6:21 am #

X

You can cut the top/bottom off a pre-trained model, add on your new dataset.



SAM September 28, 2018 at 12:43 pm #

Hi Sir I am seeking for a help if you do not mind

I am using transfer-learning to classify a dataset of images with good tutorial on tensorflow website

https://www.tensorflow.org/hub/tutorials/image_retraining

I run the code on tensorflow virtual environment on my machine and it works like they want. what I am looking for is what should I add or change in the code "retrain.py" to show the confusion matrix with python and Thanks.

Your Start in Machine Learning

You can master applied Machine Learning without the math or fancy degree.

Find out how in this *free* and *practical* email course.

Email Address

START MY EMAIL COURSE



Jason Brownlee September 28, 2018 at 3:01 pm #

REPLY ↩

You can learn how to implement a confusion matrix here:

<https://machinelearningmastery.com/confusion-matrix-machine-learning/>



SAM October 3, 2018 at 8:45 am #

REPLY ↩

Hi Sir

Thank youuuu So Much for Your Help and Your Kindness <3



Jason Brownlee October 3, 2018 at 4:13 pm #

REPLY ↩

No problem.

Your Start in Machine Learning



Nikronic November 22, 2018 at 11:07 pm #

REPLY ↩

Hello

Very good explanation. Thank you

There is a little problem, link to google Inception model is broken. Please update the link.



Jason Brownlee November 23, 2018 at 7:52 am #

REPLY ↩

Thanks, fixed. Here's the link:

<https://github.com/tensorflow/models/tree/master/research/inception>



nandini December 17, 2018 at 11:18 pm #

can we apply transfer learning on RNN model using keras



Jason Brownlee December 18, 2018 at 6:02 am #

Your Start in Machine Learning

You can master applied Machine Learning **without the math or fancy degree**. Find out how in this *free* and *practical* email course.

Email Address

START MY EMAIL COURSE



Cathy January 29, 2019 at 9:23 am #

REPLY ↩

Thanks Jason for the nice work, as always!

If I understand correctly, transfer learning is largely focused in using pretrained model on a large data set to make predictions on new data. I am wondering how transferrable is pretrained model on a not so large data set, i.e., 10k entries, to similar problems? For example, I have trained a good neural network to do sentiment analysis on ~10k tweets. How confident should I be to use this model to do sentiment analysis on other problems?

Thanks in advance!



Jason Brownlee January 29, 2019 at 11:40 am #

REPLY ↩

Typically a model fit on a very large dataset is a good starting point for use on small datasets.

Going from small to small datasets – it really depends on the tasks involved. I recommend prototype some examples and evaluate performance.

Leave a Reply

Your Start in Machine Learning

Name (required)

Email (will not be published) (required)

Website

SUBMIT COMMENT

Welcome to Machine Learning Mastery!



Hi, I'm Jason Brownlee, PhD
I write tutorials to help developers (*like you*) get results.

[Read More](#)

Your Start in Machine Learning

You can master applied Machine Learning **without the math or fancy degree**.
Find out how in this *free* and *practical* email course.

Email Address

[START MY EMAIL COURSE](#)

Better Deep Learning

Faster training, less overfitting and ensembles.

[Click to Get Started Now!](#)



Your Start in Machine Learning



So, You are Working on a Machine Learning Problem...

APRIL 4, 2018



How to Develop LSTM Models for Multi-Step Time Series Forecasting of Household Power Consumption

OCTOBER 10, 2018



11 Classical Time Series Forecasting Methods in Python (Cheat Sheet)

AUGUST 6, 2018



How to Make Predictions with Keras

APRIL 9, 2018



How to Develop LSTM Models for Time Series Forecasting

NOVEMBER 14, 2018



You're Doing it Wrong. Why Machine Learning Does Not Have to

MARCH 28, 2018



How to Make Predictions with scikit-learn

APRIL 6, 2018

Your Start in Machine Learning

You can master applied Machine Learning **without the math or fancy degree.**

Find out how in this *free* and *practical* email course.

Email Address

START MY EMAIL COURSE

You might also like...

- [How to Install Python for Machine Learning](#)
- [Your First Machine Learning Project in Python](#)
- [Your First Neural Network in Python](#)
- [Your First Classifier in Weka](#)
- [Your First Time Series Forecasting Project](#)

© 2019 Machine Learning Mastery. All Rights Reserved.

[Privacy](#) | [Disclaimer](#) | [Terms](#) | [Contact](#)