SUBSTITUTE TEACHER NETWORKS: LEARNING WITH ALMOST NO SUPERVISION

Samuel Albanie[*]
British Institute of Learning, Yearning and Discerning
Shelfanger, UK

James Thewlis[*]
National Academy of Pseudosciences
Valencia, Spain

João F. Henriques[*]
Oxford Analytika
Coimbra, Portugal

ABSTRACT

Learning through experience is time-consuming, inefficient and often bad for your cortisol levels. To address this problem, a number of recently proposed teacher-student methods have demonstrated the benefits of *private tuition*, in which a single model learns from an ensemble of more experienced tutors. Unfortunately, the cost of such supervision restricts good representations to a privileged minority. Unsupervised learning can be used to lower tuition fees, but runs the risk of producing networks that require *extracurriculum learning* to strengthen their CVs and create their own LinkedIn profiles[1]. Inspired by the logo on a promotional stress ball at a local recruitment fair, we make the following three contributions. First, we propose a novel *almost no supervision* training algorithm that is effective, yet highly scalable in the number of student networks being supervised, ensuring that education remains affordable. Second, we demonstrate our approach on a typical use case: learning to bake, developing a method that tastily surpasses the current state of the art. Finally, we provide a rigorous quantitive analysis of our method, proving that we have access to a calculator[2]. Our work calls into question the long-held dogma that life is the best teacher.

> Give a student a fish and you feed them for a day, teach a student to gatecrash seminars and you feed them until the day they move to Google.
>
> *Andrew Ng, 2012*

1   INTRODUCTION

Since time immemorial, learning has been the foundation of human culture, allowing us to trick other animals into being our food. The importance of teaching in ancient times was exemplified by Pythagoras, who upon discovering an interesting fact about triangles, soon began teaching it to his followers, together with some rather helpful dietary advice about the benefits of avoiding

---

[*]Authors listed in order of the number of guinea pigs they have successfully taught to play competitive bridge. Ties are broken geographically. This submission is a *post-print* (an update to the conference edition).

[1]Empirically, we have observed that this is extremely important for their job prospects, since it allows them to form new connexions.

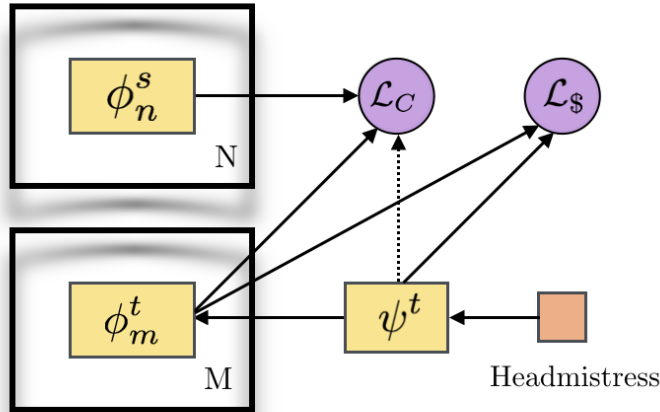[2]For all experimental results reported in this paper, we used a Casio FX-83GTPLUS-SB-UT.

Figure 1: We introduce *Substitute Teacher Networks*, a financially prudent approach to student network education. Here, $\mathcal{L}_C$ dnotes the classroom distillation loss and $\mathcal{L}_\$$ denotes the total cost of teacher remuneration. The student, task-specific teacher and substitute teacher networks are denoted by ϕ_n^s, ϕ_m^t and ψ^t respectively (see Sec. 3 for details). Note the use of drop-shadow plate notation, which indicates the direction of the nearest light source.

beans (Philolaus of Croton, 421 BC). Despite this auspicious start, his significant advances on triangles and beans reached a limited audience, largely as a consequence of his policy of forbidding his students from publishing pre-prints on arXiv and sharing source code. He was followed, fortuitously, by the more open-minded Aristotle, who founded the open-access publishing movement and made numerous contributions to Wikipedia beyond the triangle and bean pages, with over 90% of his contributions made on the page about Aristotle.

Nowadays, we are attempting to pass on this hard-won knowledge to our species' offspring, the machines (Timberlake, 2028; JT-9000, 2029)[3], who will hopefully keep us around to help with house chores. Several prominent figures of our time (some of whom know their CIFAR-10 from their CIFAR-100) have expressed their reservations with this approach, but really, what can possibly go wrong?[4] Moreover, several prominent figures in our paper say otherwise (Fig. 1, Fig. 2).

The objective of this work is therefore to concurrently increase the knowledge and reduce the ignorance, or more precisely *gnorance*[5] of student artificial neural networks, and to do so in a fiscally responsible manner given a fixed teaching budget. Our approach is based on *machine learning*, a recently trending topic on Twitter. Formally, define a collection of teachers $\{T_e\}$ to be a set of highly educated functions which map frustrating life experiences (typically real) into extremely unfair exam questions in an examination space (typically complex, but often purely imaginary). Further, define a collection of students $\{S_t\}$ as a set of debt ridden neural networks, initialised to be noisy and random. Pioneering early educational work by Bucilua et al. (2006) demonstrated that by pursuing a carefully selected syllabus, an arbitrary student S_t could improve his/her performance with M highly experienced, specialist teachers - an approach often referred to as the *private tuition* learning paradigm. While effective in certain settings, this approach *does not scale*. More specifically, this algorithm scales in cost as $\mathcal{O}(\$MNK)$, where N is the number of students, M is the number of private tutors per student and $\$K$ is the price the bastards charge per hour. Our key observation is that there is a cheaper route to ignorance reduction, which we detail in Sec. 3.

---

[3]The work of these esteemed scholars indicates the imminent arrival of general Artificial Intelligence. Their methodology consists of advising haters, who might be inclined to say that it is fake, to take note that it is in fact so real. The current authors, not having a hateful disposition, take these claims at face value.

[4]This question is rhetorical, and should be safe to ignore until the Ampere release.

[5]The etymology of network *gnorance* is a long and interesting one. Phonetic experts will know that the *g* is silent (cf. the silent *k* in knowledge), while legal experts will be aware that the preceding *i* is conventionally dropped to avoid costly legal battles with the widely feared litigation team of Apple Inc.

2 RELATED WORK

> You take the blue pill—the story ends, you
> wake up in your bed and believe whatever
> you want to believe. You take the red
> pill—you stay in Wonderland, and I show
> you how deep the ResNets go.
>
> *Kaiming He, 2015*

Several approaches have been proposed to improve teaching quality. Work by noted entomologists Dean, Hinton and Vinyals illustrated the benefits of comfortable warmth in enabling students to better extract information from their teachers (Hinton et al., 2015). In more detail, they advocated adjusting the value of T in the softmax distribution:

$$p_i = \frac{\exp(x_i/T)}{\sum_j \exp(x_j/T)},\tag{1}$$

where T denotes the wattage of the classroom storage heater. However, Rusu et al. (2015), who rigorously evaluated a range of thermostat options when teaching games, found that turning *down* the temperature to a level termed "Scottish"[6], leads to better breakout strategies from students who would otherwise struggle with their Q-values.

Alternative, thespian-inspired approaches, most notably by Parisotto et al. (2016), attempted to teach students not only *which* action should be performed, but also *why* (see also Gupta et al. (2016) for more depth). Many of the students did not want to know why, and refused to take any further drama classes. This is a surprising example of students themselves encouraging the pernicious practice of *teaching to the test-set*. Recent illuminating work by leading light and best-dressed Thessalonian (Belagiannis et al., 2018) has shown that these kind of explanations may be more effective in an adversarial environment. More radical approaches have advocated the use of alcohol in the classroom, something that we do not condone directly, although we think it shows the right kind of attitude to innovation in education (Crowley et al., 2017).

Importantly, all of the methods discussed above represent *financially unsustainable* ways to extract knowledge that is already in the computer (Zoolander, 2004). Differently from these works, we focus on the *quantity*, rather than the *quality* of our teaching method. Perhaps the method most closely related to ours was recently proposed by Schmitt et al. (2018). In this creative work, the authors suggest kickstarting a student's education with intensive tuition in their early years, before letting them roam free once they feel confident enough to take control of their own learning (the method thus consists of separate stages, separated by puberty). While clearly an advance on the expensive nanny-state approach advocated by previous work, we question the wisdom of handing over complete control to the student. We take a more responsible approach, allowing us to reduce costs while still maintaining an appropriate level of oversight.

A different line of work on learning has pursued punchy three-verb algorithms, popularised by the seminal "attend, infer, repeat" (Eslami et al., 2016) approach. Attendance is a prerequisite for our model, and cases of truancy will be reported to the headmistress (see Fig 1). Only particularly badly behaved student networks will be required to repeatedly "look, listen and learn" (Arandjelovic & Zisserman, 2017) the lecture course as many times as it takes until they can "ask, attend and answer" (Xu & Saenko, 2016) difficult questions on the topic. These works pursue a longstanding research problem: how to help models really find themselves so that they can "eat, pray and love" (Gilbert, 2009).

We note that we are not the first to consider the obvious and appropriate role of capitalism in the teaching domain. A notable trend in the commoditisation of education is the use of MOOCs (Massive Open Online Courses) by large internet companies. They routinely train thousands of student networks in parallel with different hyperparameters, then keep only the top-performer of the class (Snoek et al., 2012; Li et al., 2016). However, we consider such practices to be wasteful and are totally not jealous at all of their resources.

---

[6]For readers who have not visited the beautiful highlands, this is approximately 3 kelvins.

A number of pioneering ideas in scalable learning under budget constraints were sensitively investigated several years ago by Maturana & Fouhey (2013). We differentiate ourselves from their approach by allowing several years to pass before repeating their findings. Inspired by the concurrent ground- and bibtex-breaking self-citing work by legendary darkweb programmers and masters of colourful husbandry (Redmon & Farhadi, 2018), we now attempt to cite a future paper, from which we shall cite the current paper (Albanie et al., 2019). This represents an ambitious attempt to send Google Scholar into an infinite depth recursion, thereby increasing our academic credibility and assuredly landing us lucrative pension schemes.

2.1 UNRELATED WORK

- A letter to the citizens of Pennsylvania on the necessity of promoting agriculture, manufactures, and the useful arts. George Logan, 1800

- Claude Debussy—The Complete Works. Warner Music Group. 2017

- Article IV Consultation—Staff Report; Public Information Notice on the Executive Board Discussion; and Statement by the Executive Director for the Republic of Uzbekistan. IMF, 2008

- A treatise on the culture of peach trees. To which is added, a treatise on the management of bees; and the improved treatment of them. Thomas Wildman. 1768

- Generative unadversarial learning (Albanie et al., 2017).

3 SUBSTITUTE TEACHER NETWORKS

We consider the scenario in which we have a set of trained, subject-specific teachers $\{\phi_m^t\}_{m=1}^M$ available for hire and a collection of student networks $\{\phi_n^s\}_{n=1}^N$, into whom we wish to distill the knowledge of the teachers. Moreover, assume that each teacher ϕ_m^t demands a certain wage, p_m. By inaccurately plagiarising ideas from the work listed in Sec. 2, we can formulate our learning objective as follows:

$$\mathcal{L} = \underbrace{\frac{1}{M}\sum_{m=1}^M \mathcal{A}_{KL}(\{\phi_n^s\}_{n=1}^N||\phi_m^t)}_{\mathcal{L}_C} + \underbrace{\lambda \sum_{m=1}^M p_m}_{\mathcal{L}_\$} \qquad (2)$$

where $\mathcal{A}_{KL}$ represents the average KL-divergence between the set of student networks and each desired task-specific teacher distribution ϕ_m^t on a set of textbook example questions and λ denotes the current value of the US Dollar in dollars (in this work, we set λ to one). By carefully examining the $\mathcal{L}_\$$ term in this loss, our key observation is that **teaching students with a large number of teachers is expensive**. Indeed, we note that all prior work has extravagantly operated in the financially unsustainable educational regime where $M >> N$.

Building on this insight, our first contribution is to introduce a novel substitute network ψ^t, which does not require formal task-specific training beyond learning to set up a screen at the front of the classroom and repeatedly play the cinematographic classic *Amélie* with (useful for any task except learning French) or without (useful for the task of learning French) subtitles. Note that the substitute teacher ψ^t provides *almost no supervisory signal*, but prevents the students from eating their textbooks or playing with the fire extinguisher. Importantly, given a wide-screen monitor and a sufficiently spacious classroom, we can replace several task-specific teachers with a single network ψ_t. The substitution of ψ_t for one or more ϕ_m^t is mediated by a headmistress gating mechanism, operating under a given set of budget constraints (see Fig. 1). In practice, we found it most effective to implement ψ_t as a Recursive Neural Network, which is defined to be the composition of a number of computational layers, and a Recursive Neural Network. In keeping with the cost-cutting focus, we carefully analysed the gradients available on the market for the $\mathcal{L}_C$ component of the loss, and after extensive research decided to use Synthetic Gradients (Jaderberg et al., 2016), which are significantly cheaper than Natural Gradients (Amari, 1998). Our resulting cost function $\mathcal{L}$, which forms the target of minimisation, is best expressed in BTC (see Fig. 2).
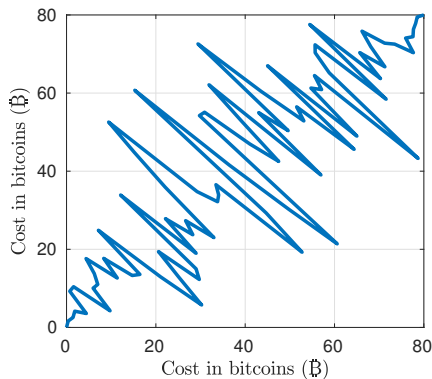
Figure 2: Expressing the cost function $\mathcal{L}$ in bitcoins makes it significantly more volatile, yet it was instrumental in attracting venture capital for our Smart Education startup.

Always driven to innovate, our second contribution is to improve upon the fiduciary example set by Enron (Sims & Brinkmann, 2003), and pay the teachers in *compound* options. These are options on options on the underlying company stock. Doing so allows us to purchase the options through a separate and opaque holding company, who then supply us with the compound options in return for a premium. The net change to our balance sheet is a fractional addition to the "Operating Activities" outgoings, and perhaps an intangible reduction in "goodwill" of the "it's not cricket" variety. The substitute teachers receive precious asymmetrical-upside however, and while a tail event would rather ruin the day, we are glass-half-full pragmatists so see no real downside. This approach draws inspiration from the Reinforcement Learning literature, which points to options as an effective extension of the payment-space (Sutton et al., 1999), especially when combined with heavy discounting.

For any given task, the high cost of private tuition severely limits the number of students that can be trained by competing methods. However, such are the scale of the cost savings that can be made with our approach that it is possible to run numerous repeats of the learning procedure. We are therefore able to formulate our educational process as a highly scalable statistical process which we call the Latent Substitute Teacher Allocation Process (LSTAP). The LSTAP is a collection of random Latent Substitute Teacher Allocations, indexed by any arbitrary input set. We state without proof a new theorem we coin the "strong" Kolmogorov extension theorem, an extension of the standard Kolmogorov[7] extension theorem (Øksendal, 2014). The strong variant allows the definition of such a potentially infinite dimensional joint allocation by ensuring that there exists a collection of probability measures which are not just consistent but *identical* with respect to any arbitrary finite cardinality Borel set of LSTAP marginals. Importantly, the LSTAP allows students to learn at multiple locations in four dimensional space-time[8].

4   EXPERIMENTS

> If you don't know how to explain MNIST to your LeNet, then you don't really understand digits.
>
> *Albert Einstein*

We now rigorously evaluate the efficacy of Substitute Teacher Networks. Traditional approaches have often gone by the mantra that it takes a village to raise a child. We attempted to use a village to train our student networks, but found it to be an expensive use of parish resources, and instead opted for the NVIDIA GTX 1080 Ti ProGamer-RGB. Installed under a desk in the office, this setup provided warmth during the cold winter months.

---

[7]While our introduction of the LSTAP may seem questionable, we have found empirically that two Kolmogorov mentions suffice to convince reviewer 2 that our method is rigorous.

[8]We follow Stephen Wolfram's definition of spacetime Wolfram (2015), and not the standard definition.

TURING TEST RESULTS

	ResNet-50	Q-Network	Neural Turing Machine
Unsupervised	C	D	
Knowledge-Distillation	B	C	F-, see me after class
Cross Modal-Distillation	A	C	
Substitute Teacher Networks (ours)	A+ ☺	B	D

Figure 3: Results for the test class of 2018. We include the Neural Turing Machine as a superficially-related baseline.

We compare the performance of a collection of student networks trained with our method to previous work that rely on *private tuition*. For a fair comparison, all experiments are performed in "library-mode", since high noise levels tend to stop concentration gradients in student networks, and learning stalls. To allow for diversity of thought amongst the students, we do not apply any of the "normalisation" practices that have become prevalent in recent research (e.g. Ioffe & Szegedy (2015); Ba et al. (2016); Ulyanov et al. (2016); Wu & He (2018)).

All students were trained in two stages, separated by lunch. We started with a simple toy problem, but the range of action figures available on the market supplied scarce mental nourishment for our hungry networks. We then moved on to pre-school level assessments, and we found that they can correctly classify most of the 10,000 digits, except for that atrocious 4 that really looks like a 9. We observed that networks trained using our method experience a much lower DropOut rate than their privately tutored contemporaries. Some researchers set a DropOut rate of 50%, which we feel is unnecessarily harsh on the student networks[9].

We finally transitioned to a more serious training regime. After months of intensive training using our trusty NVIDIA desk-warmer, which we were able to compress down to two days using montage techniques and an 80's cassette of Survivor's "Eye of the Tiger", each cohort of student networks were ready for action. The only appropriate challenge for such well-trained networks, who eat all well-formed digits for breakfast, was to pass the Turing test. We thus embarked on a journey to find out whether this test was even appropriate.

The Chinese Room argument, proposed by Searle (1980) in his landmark paper about the philosophy of AI, provides a counterpoint. It is claimed that an appropriately monolingual person in a room, equipped with paper, pencil, and a rulebook on how to respond politely to any written question in Chinese (by mapping appropriate input and output symbols), would appear from the outside to speak Chinese, while the person in the room would not actually understand the language. However, over the course of numerous trips to a delicious nearby restaurant, we gradually discovered that the contents of our dessert fortune cookies could be strung together, quite naturally, to form a message: "Flattery will go far tonight. He who throws dirt is losing ground. Never forget a friend. Especially if he owes you. Remember to backup your data. P.S. I'm stuck in a room, writing fortune cookie messages". Since we may conclude that such a message could only be constructed by an agent with an awareness of their surroundings *and* a grasp of the language, it follows that Searle's argument does not hold water. Having resolved all philosophical and teleological impediments, we then turned to the application of the actual Turing tests.

Analysing the results in Table 3, we see that only the ResNet-50 got a smiley face. The Q-network's low performance is obviously caused by the fact that it plays too many Atari games. However, we note that it could improve by spending less time on the Q's and more time on the A's. The Neural Turing Machine (NTM) had an abysmal score, which we later understood was because it focused on an entirely different Turing concept. The Q-network and the NTM disrupted the test by starting to play Battleship, and the Neural Turing Machine won.[10]

---

[9]This technique, often referred to in the business management literature as Rank-and-Yank (Amazon), may be of limited effectiveness in the classroom.

[10]We attribute this to its ability at decoding enemy's submarine transmissions.

Figure 4: Several cakes of importance for current research (deeper is better (Ardalani et al., 2018)). From left to right: 1) Yann LeCun's cake, 2) Pieter Abbeel's cake, 3) Our cake. Note the abundance of layers in the latter.

5 APPLICATION: LEARNING TO BAKE

As promised in the mouth watering abstract and yet undelivered by the paper so far, we now demonstrate the utility of our method by applying it one of the hottest subjects of contemporary machine learning: learning to bake. We selected this task, not because we care about the state-of-the-tart, but in a blatant effort to improve our ratings with the sweet-toothed researcher demographic[11].

We compare our method with a number of competitive cakes that were recently proposed at high-end cooking workshops (LeCun, 2016; Abbeel, 2017) via a direct bake-off, depicted in Fig. 4. While previous authors have focused on cherry-count, we show that better results can be achieved with more layers, without resorting to cherry-picking. The layer cake produced by our student networks consists of more layers than any previous cake (Fig. 4-3), showcasing the depth of our work[12]. Note that Jégou et al. (2017) claim to achieve a 100-layer tiramisu, which is technically a cake, but a direct comparison would be unfair because it would undermine our main point.

We would like to dive deep into the technical details of our novel use of the No Free Lunch Theorem, Indian Buffet Processes and a Slow-Mixing Markov Blender, but we feel that increasingly thin culinary analogies are part of what's wrong with contemporary Machine Learning (Rahimi, 2017).

6 CONCLUSION

This work has shown that it is possible to achieve cost efficient tuition through judicious use of Substitute Teacher Networks. A major finding of this work, found during cake consumption, is that current networks have a Long Short-Term Memory, but they also have a Short Long-Term Memory. The permutations of Short-Short and Long-Long are left for future work, possibly in the short-term, but probably in the long-term.

ACKNOWLEDGEMENTS

REFERENCES

Abbeel, Pieter. Keynote Address: Deep Learning for Robotics. 2017.

Albanie, Samuel, Ehrhardt, Sébastien, and Henriques, João F. Stopping gan violence: Generative unadversarial networks. *Proceedings of the 11th ACH SIGBOVIK Special Interest Group on Harry Quechua Bovik.*, 2017.

---

[11]The application domain was recommended by our marketing team, who told us that everyone likes cake.

[12]The code for this experiment is available at: `https://github.com/albanie/SIGBOVIK18-STNs`

Albanie, Samuel, Ehrhardt, Sébastien, Thewlis, James, and Henriques, João F. Defeating google scholar with citations into the future. *Proceedings of the 13th ACH SIGBOVIK Special Interest Group on Harry Quechua Bovik.*, 2019.

Amari, Shun-Ichi. Natural gradient works efficiently in learning. *Neural computation*, 10(2):251–276, 1998.

Amazon. Details redacted due to active NDA clause.

Arandjelovic, Relja and Zisserman, Andrew. Look, listen and learn. In *2017 IEEE International Conference on Computer Vision (ICCV)*, pp. 609–617. IEEE, 2017.

Ardalani, Newsha, Hestness, Joel, and Diamos, Gregory. Have a larger cake and eat it faster too: A guideline to train larger models faster. *SysML*, 2018.

Ba, Jimmy Lei, Kiros, Jamie Ryan, and Hinton, Geoffrey E. Layer normalization. *arXiv preprint arXiv:1607.06450*, 2016.

Belagiannis, Vasileios, Farshad, Azade, and Galasso, Fabio. Adversarial network compression. *arXiv preprint arXiv:1803.10750*, 2018.

Bucilua, Cristian, Caruana, Rich, and Niculescu-Mizil, Alexandru. Model compression. In *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 535–541. ACM, 2006.

Crowley, Elliot J, Gray, Gavin, and Storkey, Amos. Moonshine: Distilling with cheap convolutions. *arXiv preprint arXiv:1711.02613*, 2017.

Eslami, SM Ali, Heess, Nicolas, Weber, Theophane, Tassa, Yuval, Szepesvari, David, Hinton, Geoffrey E, et al. Attend, infer, repeat: Fast scene understanding with generative models. In *Advances in Neural Information Processing Systems*, pp. 3225–3233, 2016.

Gilbert, Elizabeth. *Eat pray love: One woman's search for everything*. Bloomsbury Publishing, 2009.

Gupta, Saurabh, Hoffman, Judy, and Malik, Jitendra. Cross modal distillation for supervision transfer. In *Computer Vision and Pattern Recognition (CVPR), 2016 IEEE Conference on*, pp. 2827–2836. IEEE, 2016.

Hinton, Geoffrey, Vinyals, Oriol, and Dean, Jeff. Distilling the knowledge in a neural network. In *Neural Information Processing Systems, conference on*, 2015.

Ioffe, Sergey and Szegedy, Christian. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167*, 2015.

Jaderberg, Max, Czarnecki, Wojciech Marian, Osindero, Simon, Vinyals, Oriol, Graves, Alex, and Kavukcuoglu, Koray. Decoupled neural interfaces using synthetic gradients. *arXiv preprint arXiv:1608.05343*, 2016.

Jégou, Simon, Drozdzal, Michal, Vazquez, David, Romero, Adriana, and Bengio, Yoshua. The one hundred layers tiramisu: Fully convolutional densenets for semantic segmentation. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2017 IEEE Conference on*, pp. 1175–1183. IEEE, 2017.

JT-9000. How I learned to stop worrying and love the machines (Official Music Video). In *British Machine Vision Conference (BMVC)*, 2029. West Butterwick, just 7 miles from Scunthorpe, England.

LeCun, Yann. Keynote Address: Predictive Learning. 2016.

Li, Lisha, Jamieson, Kevin, DeSalvo, Giulia, Rostamizadeh, Afshin, and Talwalkar, Ameet. Hyperband: A novel bandit-based approach to hyperparameter optimization. *arXiv preprint arXiv:1603.06560*, 2016.

Maturana, Daniel and Fouhey, David. You Only Learn Once - A Stochastically Weighted AGGRegation approach to online regret minimization. In *Proceedings of the 7th ACH SIGBOVIK Special Interest Group on Harry Quechua Bovik.*, 2013.

Øksendal, Bernt. *Stochastic Differential Equations: An Introduction with Applications (Universitext)*. Springer, 6th edition, January 2014. ISBN 3540047581. URL `http://www.amazon.com/exec/obidos/redirect?tag=citeulike07-20&path=ASIN/3540047581`.

Parisotto, Emilio, Ba, Jimmy, and Salakhutdinov, Ruslan. Actor-mimic: Deep multitask and transfer reinforcement learning. In *ICLR*, 2016.

Philolaus of Croton. How to bake the perfect croton. *Greek Journal of Fine, Fine Bean-Free Cuisine*, 421 BC.

Rahimi, Ali. Test of Time Award Ceremony. 2017.

Redmon, Joseph and Farhadi, Ali. Yolov3: An incremental improvement. 2018.

Rusu, Andrei A, Colmenarejo, Sergio Gomez, Gulcehre, Caglar, Desjardins, Guillaume, Kirkpatrick, James, Pascanu, Razvan, Mnih, Volodymyr, Kavukcuoglu, Koray, and Hadsell, Raia. Policy distillation. *arXiv preprint arXiv:1511.06295*, 2015.

Schmitt, Simon, Hudson, Jonathan J, Zidek, Augustin, Osindero, Simon, Doersch, Carl, Czarnecki, Wojciech M, Leibo, Joel Z, Kuttler, Heinrich, Zisserman, Andrew, Simonyan, Karen, et al. Kickstarting deep reinforcement learning. *arXiv preprint arXiv:1803.03835*, 2018.

Searle, John R. Minds, brains, and programs. *Behavioral and brain sciences*, 3(3):417–424, 1980.

Sims, Ronald R and Brinkmann, Johannes. Enron ethics (or: culture matters more than codes). *Journal of Business ethics*, 45(3):243–256, 2003.

Snoek, Jasper, Larochelle, Hugo, and Adams, Ryan P. Practical bayesian optimization of machine learning algorithms. In *Advances in neural information processing systems*, pp. 2951–2959, 2012.

Sutton, Richard S, Precup, Doina, and Singh, Satinder. Between mdps and semi-mdps: A framework for temporal abstraction in reinforcement learning. *Artificial intelligence*, 112(1-2):181–211, 1999.

Timberlake, Justin. Filthy (Official Music Video). In *Pan-Asian Deep Learning Conference*, 2028. Kuala Lumpur, Malaysia.

Ulyanov, Dmitry, Vedaldi, Andrea, and Lempitsky, Victor. Instance normalization: The missing ingredient for fast stylization. *arXiv preprint arXiv:1607.08022*, 2016.

Wolfram, Stephen. What is spacetime really?, 2015. `http://blog.stephenwolfram.com/2015/12/what-is-spacetime-really/`.

Wu, Yuxin and He, Kaiming. Group normalization. *arXiv preprint arXiv:1803.08494*, 2018.

Xu, Huijuan and Saenko, Kate. Ask, attend and answer: Exploring question-guided spatial attention for visual question answering. In *European Conference on Computer Vision*, pp. 451–466. Springer, 2016.

Zoolander, Derek et. al. *Zoolander*. Paramount Pictures, 2004.