

**Chapter 1 : A Look at the Growth of Cloud Services**

*Free draft copy of Andrew Ng's book - Machine Learning Yearning!*

Expand to view full transcript Kirill Eremenko: Welcome to the Super Data Science Podcast. My name is Kirill Eremenko, data science coach and lifestyle entrepreneur. Each week, we bring you inspiring people and ideas to help you build your successful career in data science. Welcome back to the Super Data Science Podcast, ladies and gentlemen. I actually met Gregory quite a while ago. KDnuggets is one of the most popular data science resources out there. Write accurate news on data science, they provide their own articles, they conduct polls on data science, and many, many more exciting things in the space of data science. Today, we are with welcoming him on the show. A very exciting, very interesting story of how it all started, where Gregory began his journey into the space and what KDnuggets has grown into, but also we will cover off some of the more recent advances that have been happening in the space of data science that KDnuggets has been highlighting or has been participating in. All in all, a very exciting episode full of most recent technology core advancements and interesting stories on how this all came to be. Welcome, ladies and gentlemen to the Super Data Science Podcast. Gregory, welcome to the show. How are you today? We met in May, 2., what was it? No, I think May 2., yeah, This is super, super, exciting. Gregory, where are you located right now? I am in Boston, Massachusetts. Yeah, I was just about to ask that. That is your home base, Boston. Last time I was there, it was in May last year, it was surprisingly chilly. Gregory, you are the Founder and Director or President and Editor of KDnuggets, a very popular data science media outlet and news aggregator and a platform that shares research about data science. Tell us a little bit about how it all started. Where did this idea come from? Probably, I started when I was a kid, I was very fascinated by science fiction, and I loved stories about robots, especially, from Isaac Asimov and other writers like Stanislaw Lem and [inaudible I was always curious about the idea of AI, and this probably motivated me to learn computers when they first year appeared. In my first year in college when computers were still programmed with punch cards, I remember spending several weeks of my free time in the summer, writing a program to play battleships, which was still a very advanced program for that period. And then I used APL. That was a special language developed by IBM. You can think of it as like R but with Greek letters. After spending several weeks programming it, I played one game and I was very soundly defeated by my own program. I think as a result, I become much more interested in creating programs than playing them. I think the idea was a self-organizing database system that automatically selects different indices and does something intelligent. Then I worked at GTE as a researcher. GTE was a large telephone company in United States. Now, it is part of Verizon, which is even a larger telecom company. I remember around or so, I attended a workshop, which was called Expert Database System. That was a very interesting name, but the concept was very fuzzy, and the workshop paper and talks were all over the place. I thought we could focus on something more clearly defined, analyzing databases and finding interesting patterns. In one of our projects that we did on applying some intelligent to figure out databases, and I discovered that a particular query would run 10, faster if we knew that there was a particular rule, that kind of functional constraint that always existed. There were some over-supplication. Can you find some useful rules in databases? At that time, a popular term was data mining. It went from data fishing and data dredging, which were bad times, and data mining became second popular term. Now, the popular term is data science or maybe until last year. That was the first workshop back in , which attracted, I think, about 70 people including several leading researchers. Then I organized a couple of more workshop, and later in , one of my best ideas was to stop doing it myself and to recruit Usama Fayyad, who was then just a fresh PhD from Ann Arbor. Now, I can stand back after many years of organizing [inaudible That was kind of one track of my activity. How did I get to where I am? The first issue, which is still online, went to, I think, about 50 people, who attended that workshop. If I try to talk to my people you realize that as a data scientist at heart, I just tried to select a few interesting things to write about or select things on the web that we can publish. I guess that was a second track in my career. In , which was still very part of the Dot-com Rush I was a chief scientist, and managed a small team of perhaps about 10 people. Then around , our smaller startup was bought by a big

startup. I left it I think maybe couple of months before that stock went all the way to zero. I was self-employed since about , mainly publishing KDnuggets and doing consulting and data mining. I think one interesting question for all the younger people listening is synergy. In each one of those activities was in some way helping the other. And probably doing other things. I guess probably, helpful suggestion for young people that try to do interesting things is to think is there a synergy with this activity with some other [inaudible The very synergy, it generally helps you to succeed. Just to finish in this, in the last few years, I think, maybe writing the big data and data science, which KDnuggets became so popular that I stopped that I stopped consulting them. Now, I only publish KDnuggets, and we have another excellent full-time idea [inaudible We have several interns based in London and other places. KDnuggets is global in its reach. That is truly astonishing numbers. You mentioned that you select those blog posts. How many blog posts do you publish on KDnuggets? How frequently do they come out? Well, we publish every weekday, and we try to select maybe two or three interesting blog posts a day. Now, we get a lot of submissions. Occasionally, myself and Matthew [May Already, that makes it 10 or 15 or more per week. How do you find the time to go through all of them? You probably get a ton of submissions sent to you. How many submissions do you get, just out of curiosity? I assume our readers already know, but we would publish something that explains how to create a pipeline in Python or some ideas how to use Python [inaudible Yeah, this is the 19th such poll. Now, the latest poll is out to show that there is kind of a clear ecosystem emerging around Python, Spark, Anaconda and TensorFlow. Python seems to have more significant [inaudible There are some other interesting observations that your readers can see on KDnuggets. Is it just like on the main page of the blog or is there a specific page for all these insights? Well, on the main menu, we have a section called top stories, and if you scroll there, then you will find more interesting things. Wow, this is really cool. Do you mind sharing a little bit on that with us? There was still a lot of unsatisfied ones, but on average, I think there was a significant difference between the job satisfaction for this profession, machine learning engineer, and the second and third place, which were researcher and data scientist. Data scientist is still the most common job title. I see that on the web and [inaudible I guess a difference I would describe as machine learning engineer is building machine learning systems, probably they now use deep learning, and data scientist perhaps do more work on analyzing and then trying to understand what is happening with companies, not necessarily building production systems. Thank you for that. All right, so I wanted to ask you a couple of questions. What is a recent win that you share with us? It was a system for analysis and summarization of key changes in large databases, and we applied it to healthcare data. Healthcare in United States is a scandal and also very, very expensive. I think we spend here twice as much as other industrialized countries.

## Chapter 2 : Free New Book by Andrew Ng: Machine Learning Yearning - Data Science Central

*Machine Learning Yearning aims at an engineer or a technical product manager for a team building a machine learning solution. The book features terse chapters that succinctly explain how to think about the high-level decisions associated with a machine learning product.*

He is writing a book, "Machine Learning Yearning" you can get a free draft copy , to teach you how to structure Machine Learning projects. Andrew writes This book is focused not on teaching you ML algorithms, but on how to make ML algorithms work. Some technical AI classes will give you a hammer; this book teaches you how to use the hammer. If you aspire to be a technical leader in AI and want to learn how to set direction for your team, this book will help. We have read the draft, and selected 7 most interesting and useful suggestions from the book: Optimizing and satisficing metrics Rather than use a single formula or metric to evaluate an algorithm, you should consider utilising multiple evaluation metrics. This can be an extremely efficient and easy way of evaluating your algorithm. An initial one-week target is set; it is better to come up with something not perfect and get going quickly, rather than overthink this stage. Just do it and let your team know the new direction you are heading in. Machine Learning is an iterative process: Start off with an idea Implement the idea in code Carry out an experiment to conclude how well the idea worked The faster you can go around this loop, the quicker progress will be made. Build your first system quickly and then iterate As mentioned in point 3, building a machine learning algorithm is an iterative process. Ng dedicates a chapter to explaining the benefits of building a first system quickly and going from there: Instead, build and train a basic system quickly—perhaps in just a few days. Included is an analysis of each image, why it failed and any additional comments that may help upon future reflection. On completion you can see which ideas will eliminate more errors and therefore which should be pursued. If you suspect that a fraction of the errors are due to this, add an additional category in your spreadsheet: Upon completion you can then consider if it is worth the time to fix these. Ng gives two possible scenarios to judge if it is worth fixing these: Overall accuracy on dev set. There is no harm in manually fixing the mislabeled images in the dev set, but it is not crucial to do so: It is now worthwhile to improve the quality of the labels in the dev set. For a project on speech recognition, in which you would be listening to audio clips, perhaps you would call this set an Ear dev set instead. The Eyeball dev set therefore has examples, of which we would expect our algorithm to misclassify about You can use the Blackbox dev set to evaluate classifiers automatically by measuring their error rates. You can also use it to select among algorithms or tune hyperparameters. However, you should avoid looking at it with your eyes.

## Chapter 3 : 7 Useful Suggestions from Andrew Ng "Machine Learning Yearning"

*Machine Learning Yearning is a calendriredelascience.com project. The subtitle of the book is Technical strategy for AI engineers in the era of deep learning. Content of the book.*

Todd Lyle May 13, Source: Lyle discusses cloud computing and its future. Everyone is talking about the cloud " not just what it is, but also how it works. But what are we doing with all of this conversation? Mobility for both stakeholders strategic and end users tactical is changing how we accomplish almost everything. Stay ahead of the curve with Techopedia! Join nearly , subscribers who receive actionable tech insights from Techopedia. The combination of tools and services you use. The versatility of the public cloud is similar to first-world utilities. This is a composition of public cloud offerings, such as email and off-site backup , while continuing to use your on-site or private cloud-based productivity resources. These companies will install your system, maintain it and act as your managed services provider. The Human Element and Cloud Technology The irony of this coming-of-age story is that the older demographics are yearning to learn more. While boomers embrace technology as a second language, Millennials continue to struggle with socializing the old-fashion way " in person. Here lays the new challenges of addressing the soft side of technology, you and me: Today we are well into and the noise in the cloud space is deafening. Enterprise-level cloud services organizations are full steam ahead with their billion-dollar budgets and a thirst for market share. The federal government in the United States is more than capable of implementing a 21st century shared-services model, yet it is encumbered by old-school thinking, compliancy and outdated purchasing models. Education and resultant legislation is required for related change. At this juncture, at a minimum, the U. Adoption by small and medium businesses is still a few years out due to a lack of simplicity in adoption in cloud services and acute security concerns. This is a huge cultural shift for both organizations and shows that all directions point towards the cloud. The consumers are the ones who win when it comes to the cost of operating an enterprise-level network via the public cloud, which is dropping to the point that it will soon be pennies on the dollar to accomplish similar computing tasks as compared to five to 10 years ago. Consumers are beginning to recognize the value of a cloud middleman. Hence, you will soon be using a cloud services brokerage to assist in cloud journey. MBA programs will soon offer a cloud program to augment their curriculum and competitiveness. Governments on all levels will embrace the concept of shared services only once it is mandated. The year will continue to find leaders looking for a one-stop-shop convenience, and this convenience will be realized with the all-encompassing proverbial desktop. Written by Todd Lyle Todd Lyle has become a leader in cloud computing with a focus on the human element. His fascination with the possibilities of information technology began when he was issued his first laptop computer while serving as a young officer in a U. Army aviation regiment in Korea. It was while working at Microsoft that Lyle became interested in the myriad ways that information technology can improve overall business productivity.

## Chapter 4 : Machine Learning Yearning by Andrew Ng

*Machine Learning Yearning, a free book that Dr. Andrew Ng is currently writing, teaches you how to structure Machine Learning projects. This book is focused not on teaching you ML algorithms, but on how to make ML algorithms work.*

The ever-evolving discipline of data science extends to almost every sector and industry imaginable, on a global scale. Gaining an astute understanding of data science as a concept as well as a practice comes with a wealth of advantages. Wondering which data science book to read? Download our free guide! For savvy data scientists, the potential to unlock this seemingly infinite ocean is colossal. Data science, also known as data-driven science, covers an incredibly broad spectrum. This interdisciplinary field of scientific methods, processes, and systems helps people extract knowledge or insights from data in a host of forms, either structured or unstructured, similar to data mining. One more time, we will go over some must-have books to add to your business intelligence bookshelf. If you would like to gain a sound comprehension of data science and take your understanding of the world to the next level, reading the best books for data science is a must. That said, these books will get your journey into the immersive world of big data off to the best possible start. Someone who has become all too aware of the machine learning and artificial intelligence craze but needs to get a grip on the subject. Driven by the acquisition and processing of complex information, machine learning is an area of data science that has emerged monumentally in recent years. Written by renowned computer scientist Andrew Ng, this gripping read not only offers an accessible introduction to machine learning and big data, but it also proves an excellent resource on collecting data, utilising the power of deep end-to-end learning, and facilitating the sharing of key insights with a machine learning system. Someone with a sound working knowledge of Python who wants to understand how to use the language to enhance their data insights. *Data Wrangling with Pandas, NumPy, and IPython* takes the reader deep into the realms of the language and its enormous potential for manipulating, processing, cleaning, and crunching data in Python. A must-read for those wrestling with Python. The CEO, Chief Digital Officer, Chief Information Officer, or business owner looking to seriously enhance their predictive analytics skills, both practically and theoretically. A *New York Times Best Seller* and for good reason *The Signal and the Noise* is a masterclass in using the power of big data analytics to make valuable predictions in an informed and potent way. Crafted by American statistician Nate Silver, a spokesperson famed for successfully predicting the US Presidential election results, this book uncovers the genuine art and science of making predictions from data. Peppared with real-world case studies, interesting examples of data prediction, and citations of epic data-based failures, this book shows the reader how to filter out the noise and hone in on the right insights to make projections that not only matter, but also ensure sustainable levels of success. The technically-minded wizard or digital tech enthusiast looking to bridge the gap between big data analytics, complex algorithms, and the way these elements will shape our future lives. Data science is largely about predictions, but a significant part of this ever-expanding discipline also boils down to sophisticated algorithms. In this thought-provoking and, in many ways, timeless work of data science prose, author, and prolific programmer Christopher Steiner explains how algorithms are increasingly being used to take on high-level pursuits that were once tackled only by human beings with niche training areas including medical diagnosis and foreign policy analysis. Once you pick it up, *Automate This: A Data Visualization Guide for Business Professionals*, by Cole Nussbaumer Knaflic Created by storytelling expert Cole Nussbaumer Knaflic, this methodical handbook is not only entertaining, but it also provides deep-rooted insights into a branch of data science that is often overlooked: In the meantime you get a hand on that book, you can have a look at our dashboard storytelling tips, to know how to tell a great story after learning how to make a dashboard in 9 easy steps. The budding data manager or data miner with a desire to make sense of information in the modern age and beyond. As far as books on data science go, this one is perhaps one of the most forward-thinking one in existence. Penned by Scott Stawski, a data management leader at Hewlett Packard, *Inflection Point* focuses on how swift changes in cloud computing, big data, mobile devices, and apps are morphing the way businesses compete. The wide-eyed, up and coming Apache Hadoop warrior with a hunger for building scalable systems from data. In one of the best

books on data science regarding processing language, Tom White takes his readers on a data-based journey to help them understand the importance of Hadoop and how, if used wisely, it can do a multitude of incredible things. These incredible things include the ability to build and manage scalable systems with Hadoop and successfully running large Hadoop clusters. The budding data scientist looking for a comprehensive introduction to the field. Straight Talk from the Frontline serves as a clear, concise, and engaging introduction to the field. Written with confidence and a clear, practical understanding of the topic, this essential guide to data science will help you hit the ground running and bestow you with the knowledge you need to thrive in this ever-growing field of expertise. When it comes to data science, there is an incredible amount to learn. In our opinion, these best data science books will help you gain the knowledge you need to embark on your long and rewarding journey towards data enlightenment.

**Chapter 5 : Machine Learning Yearning**

*1 Why Machine Learning Strategy Machine learning is the foundation of countless important applications, including web search, email anti-spam, speech recognition, product recommendations, and more.*

Machine Learning techniques can help modern businesses run much more efficiently and in a more predictable way. In this article, I am going to scratch the surface of machine learning basics and explain what the key takeaways are for business owners. In general, machine learning is a technique, where behavior of an application or algorithm changes based on past iteration. Does it mean a birth of new Skynet? This misunderstanding is created from misinterpreting the words of Elon Mask, for example. Basically, each machine learning task is constructed from the following components: Dataset of samples of data or Training Sets TS that are used as input for initial algorithm. An example for a business would be data about the number of people living in a specific city and the number of sales in that city. An initial hypothesis that we would like to improve using ML technique. Cost function, or a formula that we use to understand how our hypothesis is working, and how accurate it is. And finally, a learning algorithm. This is where the magic happens. The rise of machine learning is happening right now, because it requires huge computation power to perform multiple iterations over the same training set. The more examples you have in a training set, and the more facts for each example you collect, the more insights you can potentially get. Same thing with the average temperature in a city, for example. Adding more facts into the equation requires much more computational power, which was not possible before. As the most extreme case of this factor-based problem, think about the weather prediction, which is a classical problem for ML. Almost any weather prediction station uses HPC high performance computing clusters to analyze and compare millions of facts to get accurate weather forecast. With modern technology, specifically CUDA solutions from NVidia, now thousands of CPUs fit on a single board or even in a single chip, which opens limitless possibilities for data researchers and there is no need for supercomputers to analyze examples that contain up to facts. If your business is using ERP or CRM system, you already have data that you can use to get insights, as you already have thousands of facts to choose from. You do not need to immediately start collecting all other data that you may get from all possible sources, rather just start analyzing data that you already have. Human resources and expertise are more important than tools for ML, as most of the tools are free for a small scale of data. Hire a data scientist and provide all Datasets that you have and start gaining insights. To gain insights in specific areas, start collecting more examples in that area, i. Use hypotheses generated by machine learning to improve your procurement and forecasting. Employ hypotheses as models to monitor your business execution and limit the number of human errors. Jump into Big Data approach only when you understand what insights you can gain for your business data. Read More From DZone.

**Chapter 6 : The Top 8 Best Data Science Books You Need To Read**

*I've just read the "Takeaways" page in the end and it's enough for now. The book deals with classical machine learning and not convolutional neural networks, so I'll get to read it fully someday later when it's out of beta and I have time to study it and Bishop's Machine learning book.*

**Chapter 7 : Machine Learning Yearning**

*Machine Learning techniques can help modern businesses run much more efficiently and in a more predictable way. In this article, I am going to scratch the surface of machine learning basics and.*

**Chapter 8 : Will Artificial Intelligence Take Away Our Jobs-Sadhguru " Machine Learning**

*Machine Learning Yearning. Contribute to*

*amusi/machine-learning-yearning-cn development by creating an account on GitHub.*

**Chapter 9 : Machine Learning Yearning - calendrierdelascience.com**

*AI, Machine Learning, and Deep Learning are rapidly evolving and transforming many industries. Andrew Y. Ng is one of the leading minds in the field - he is a co-Founder of Coursera, former head of Baidu AI Group, and a former head of Google Brain. He is writing a book, "Machine Learning Yearning."*