Nic Payne
332 West Parkway Blvd
Appleton, WI 54911
(319) 389-5740

December 23, 2021

Luke & LMG Team
Linus Media Group
The Great White North

Luke and LMG folks,

Thanks you including me in the pool of applicants for the Data Science role at LMG. I've been following LTT and related channels for a few years and have been dragged deep into the water of computer building and home-labbing. The announcement of the Data Scientist position got my wife and I pretty jazzed because we're big fans of the content and what the material from LMG has enabled me to do at home and at my current job.

I think it makes fair sense to first say that I'm not necessarily looking to leave my current job but would like to see if the possibility of a remote DS freelance consultant fits into the plans at LMG (however I'm not married to the idea of staying either). I've been doing data engineering and data scientist work for 4.5 years at Caterpillar, with the last 2 years obviously being fully remote - much of that experience is captured in my resume - but one thing I've learned is that I can do this type of work for anyone from anywhere. In the last 2 years being fully remote I have been able to successfully coach other data scientists and python programmers within the Cat enterprise through internal group meetings like our "Breakfast and Supper Learn" which consists of several dozen pythonistas from around the world as well as on my team where I work alongside 4 other data scientists. I also have been able to excel at my work because the tools needed for practicing and deploying data engineering/scientist solutions aren't steel and iron that I have to hold in my hand, and they aren't even Excel or a desktop app that someone in an office can look at on my computer, but instead it's cloud computing/hosting, automated alerts/notifications from AWS, scheduled pipelines, etc. I have confidence that I can help someone meet a data need from anywhere with a stable-enough internet connection.

So why am I interested in working with/at LMG then? I like what I get to work on with Cat but a lot my excitement is related to the tools I get to use, not the data necessarily. A position with LMG would bring excitement over subject matter as well as tooling. I think working with data from gaming benchmarks, temperatures during a stress test, etc. would be an incredibly fun and fullfilling content-matter to which I can fluently apply data engineering and data science tools and solutions. I also feel like I owe you guys some value - my work from home office space is a very comfortable environment and I owe part of that to what I've learned about cameras, network, wifi and router placement, laptop cooling, and the importance of a slick LTT Northern Lights Desk Pad (ltt store dot com) to bring it all together.

I tried to outline key achievements in my resume but I'll highlight what I think the important ones are here. During my time in EXD and ICS at Cat in 2018-2019 I was tasked with designing an end to end machine learning pipeline initially targeting personnel detection for on-board safety systems but with a larger goal of cloud-native ML development across Caterpillar. The design was entirely up to me as we started with nothing. What I came up with included deploying python applications in docker which indexed new and existing data (camera/lidar/radar) on an on-prem NAS, applying a deep learning object detection based grooming and cleaning algorithm to filter out unusable data and tag good data with key observations (if a person or a machine was present in a frame mainly). The solution included uploading clean datasets to S3 to be used by other groups in Cat working on computer vision applications. The end goal was to enable the creation of deep learning based computer vision and perception systems for on-machine deployment (ie. personnel detection on an autonomous excavator).

I moved on from that position into a group where the data was not images/perception systems based but more traditional tabular business data. In my current role I design ETL pipelines using a framework called Kedro which is 100% FOSS and written in Python. I do some development on my team for custom Kedro utilities and focus a lot on the infrastrucure underlying our Kedro deployment (dockerization, CI/CD, scheduled deployment on AWS Batch). These 2 roles that I've filled at Cat have given me a lot of experience, and I'd even say some expertise, on designing data pipelines, thinking through and even designing database schemas for efficient relational data storage, and applying data science tools to that data.

I'd be ecstatic for the opportunity to interview and discuss what it might look like for me to serve LMG as a remote data scientist who can hopefully bring a unique blend of not only data science skills, but data engineering skills as well to the team.

It's probably fine,

Nicholas Payne