

## Introdução

A LTrace desenvolve soluções de Inteligência Artificial para a indústria de petróleo com o objetivo de otimizar o processo de atualização de modelos de reservatório. Utilizando tecnologias como deep learning, a LTrace foca principalmente na redução em dados e grande dimensão e geração de dados categóricos, empregando Redes Geradoras para aprimorar a eficiência e precisão dos modelos.

Redes Geradoras são tipos de modelos de aprendizado de máquina usados para gerar novos dados semelhantes aos dados de treinamento. As VAEs (Variational Autoencoders) , utilizam uma abordagem de codificação (estilo Autoencoder) variacional para aprender a distribuição dos dados e gerar novas amostras ao decodificá-las. Já as GANs (Generative Adversarial Networks), consistem em duas redes neurais (gerador e discriminador) que competem entre si para criar dados realistas. Ambos os modelos são amplamente utilizados em tarefas como geração de imagens, síntese de voz e criação de dados sintéticos.

## Objetivo

O objetivo do teste é utilizar redes neurais 2D do tipo geradoras, VAE ou GAN, para gerar novas imagens de canais a partir de um espaço latente gerado aleatoriamente.

## Dados

O conjunto de dados é formado por 50 mil imagens de tamanho 128x128 pixels que representam recortes de canais geológicos do tipo meandro. O amarelo (1) representa uma classe canal e o azul (0) representa a classe não-canal.



Os dados estão em um arquivo denominado *train\_images.h5* no buckets da oracle:

[https://grrjnyzvhu1t.objectstorage.sa-saopaulo-1.oci.customer-oci.com/p/jflQ-fCMHmA5jLi-DdqElenJTzARdTp61YIzIDpRBI531v05mmWObjM2F1CBTIVL/n/grrjnyzvhu1t/b/General\\_ltrace\\_files/o/deep\\_esmda/train\\_images.h5](https://grrjnyzvhu1t.objectstorage.sa-saopaulo-1.oci.customer-oci.com/p/jflQ-fCMHmA5jLi-DdqElenJTzARdTp61YIzIDpRBI531v05mmWObjM2F1CBTIVL/n/grrjnyzvhu1t/b/General_ltrace_files/o/deep_esmda/train_images.h5)

## Critérios

O tratamento do conjunto de dados é livre, podendo inclusive redimensionar as imagens para um tamanho menor se preferir, basta somente descrever o que foi feito.

Independente do modelo utilizado, após o treinamento, o espaço latente deve estar regularizado. Uma introdução sobre espaço latente regularizado pode ser encontrado em <https://towardsdatascience.com/understanding-variational-autoencoders-vaes-f70510919f73>

O modelo desenvolvido deve conter camadas convolucionais. Como ponto extra, seria interessante o candidato comparar os resultados de dois modelos: um modelo formado por camadas convolucionais + densas (fully connected) e outro modelo composto somente camadas convolucionais (fully convolutional). Neste último modelo, não se deve utilizar nenhuma camada densa em toda a estrutura da rede, incluindo o espaço latente.

A plataforma de machine learning pode ser a escolha do candidato, porém para que possamos auxiliar da melhor forma, recomendamos PyTorch (preferência) ou Tensorflow / Keras em python.

O código desenvolvido deve ser disponibilizado em um sistema de controle de revisão de código (Github, Bitbucket, ...).

Espera-se que após o teste, o candidato discuta as ideias que teve durante o planejamento e modelagem da rede neural e explique o que entendeu sobre a aplicação requisitada.

## Prazo

O prazo de entrega é de 7 dias a partir do recebimento. Estima-se que a estrutura básica de cada VAE ou GAN leva cerca de 1-2 horas para ser implementada.

## Envio do teste e Dúvidas

O teste (link do repositório) e qualquer dúvida podem ser enviados por e-mail para [fernando@ltrace.com.br](mailto:fernando@ltrace.com.br) e [rodrigo@ltrace.com.br](mailto:rodrigo@ltrace.com.br)