

Trabalho Classificador de Dígitos Mnist

O desenvolvimento de um sistema baseado em algoritmos de inteligência computacional para o reconhecimento de caracteres faz é utilizado em inúmeras aplicações, sendo uma área de grande interesse no cenário científico e empresarial, de modo a automatizar tarefas.

Um sistema de reconhecimento de caracteres possibilita a verificação e/ou classificação de padrões que aplicado em imagens possibilita o Reconhecimento Óptico de Caracteres (OCR). Essa tecnologia permite que uma máquina possa reconhecer automaticamente um caractere através de um mecanismo óptico.

O OCR consiste em um sistema de visão computacional que, inicialmente, obtém uma imagem digitalizada do texto a ser reconhecido, normaliza a imagem em tons de cinza e por fim realiza a segmentação dos caracteres baseados em suas bordas. Com as imagens dos caracteres segmentados é aplicado um algoritmo de inteligência artificial que o identifica e pode até mesmo reconhecer o seu significado semântico no contexto (*chatbot*).

Assim, este trabalho consiste no desenvolvimento de um sistema de inteligência computacional baseado em redes neurais artificiais (MLP) para o reconhecimento de dígitos numéricos. Para o treinamento da rede será utilizada a base de dados MNIST que consiste em 70000 caracteres digitalizados em tons de cinza em imagens de 28x28 px, portanto a parte da visão computacional já foi realizada. Assim, o banco de dados está separado em três arquivos no formato CSV (1ª coluna é o rótulo da classe) um com 60 mil entradas (dígitos) de 784 colunas (imagem de 28x28 colocada em uma única linha), 10 mil entradas e 100 entradas.

Para o treinamento e avaliação da melhor configuração pode ser utilizada qualquer biblioteca de MLP existente, sendo que devem ser avaliadas no mínimo seis topologias candidatas a livre escolha do aluno, sendo três com uma camada oculta e três com duas camadas ocultas (podem ser usado os critérios eurísticos apresentados nos slides, três equações). Para a avaliação utilizar o arquivo com 10000 imagens e uma das técnicas de *cross-validation* apresentadas.

Após definida a topologia a ser empregada realizar o treinamento com o arquivo de 60000 imagens e realizar o teste com o arquivo de 10000 imagens, obtendo a precisão e a matriz de confusão.

Com o treinamento executado obter os pesos da rede neural e desenvolver um script (python ou C) para o reconhecimento de caracteres utilizando o algoritmo de operação apresentado em aula. O script deve possuir uma função que terá com entrada um vetor de 784 colunas correspondente ao dado de entrada de uma imagem obtida pelo sistema de visão computacional e retornará o dígito previsto pelo algoritmo de inteligência computacional.

Para o teste do script de operação deve ser utilizado o arquivo de 100 imagens e obtido a precisão e matriz de confusão do algoritmo.

Com o trabalho finalizado deve ser realizado um relatório descrevendo detalhadamente o que foi executado, com os materiais e métodos utilizados, no formato artigo já utilizado. O artigo em formato PDF juntamente com os scripts desenvolvidos devem ser enviados para o e-mail fabianomora@pelotas.ifsul.edu.br até o dia **06 de junho de 2019**.

Os 3 arquivos para a execução do trabalho estão no link: https://drive.google.com/drive/folders/1lufCqEo8NbJNEPaN-07XmIQ-IT4_qOPP?usp=sharing

Qualquer dúvida entrar em contato pelo e-mail.

Bom trabalho a todos!