https://eventos.utfpr.edu.br//sei/sei2019

Robô Jardineiro Bioinspirado com reconhecimento de padrões pelo classificador k-NN

Bioinspired gardner robot with pattern recognition by k-NN classifier

Débora de Azevedo Rodrigues
debora9rodrigues@gmail.com
Universidade Tecnológica Federal do Paraná, Toledo, Paraná, Brasil

Daniel Cavalcanti Jeronymo
danielc@utfpr.edu.br
Universidade Tecnológica Federal do Paraná, Toledo, Paraná, Brasil

Juan Manoel Marinho Nascimento
juanengml@gmail.com
Universidade Tecnológica Federal do Paraná, Toledo, Paraná, Brasil

RESUMO

Neste trabalho é apresentado um sistema de irrigação inteligente onde é proposta uma solução para o problema de tomada de decisão na irrigação. A proposta constitui-se em dispositivos *IoT* que monitoram a umidade e temperatura do solo, utilizando sensores conectados a um *Arduino*. Os dados destes dispositivos alimentam um servidor com a plataforma *ThingsBoard* onde os dados são armazenados e analisados. As informações coletadas são utilizadas como entrada de um modelo pré-treinado utilizando o classificador k-NN. A saída do sistema é dada por uma regra de tomada de decisão da irrigação. Os resultados obtidos apontam a viabilidade da proposta quanto a manutenção de umidade do solo.

PALAVRAS-CHAVE: Irrigação Inteligente. Classificador k-NN. Internet das Coisas.

ABSTRACT

In this paper a smart irrigation system is presented with a proposed solution for the irrigation decision making problem. This proposal consists in IoT devices for monitoring soil humidity and temperature, using sensors connected to a Arduino. Data from these devices feeds a server running the platform ThingsBoard where data is stored and analyzed. Collected information is then used as the input for a pre-trained k-NN classifier. The system's output is given by an irrigation decision making rule. Obtained results indicate feasibility of the proposal regarding maintenance of soil humidity.

KEYWORDS: Smart Irrigation. k-NN Classifier. Internet of Things.

Recebido: 19 ago. 2019.
Aprovado: 01 out. 2019.

Direito autoral: Este trabalho está licenciado sob os termos da Licença Creative Commons-Atribuição 4.0 Internacional.

Insira aqui o selo referente ao Objetivo do Desenvolvimento Sustentável que o trabalho atende diretamente (Obrigatório).

Insira aqui o selo referente ao Objetivo do Desenvolvimento Sustentável que o trabalho atende direta ou indiretamente.

Insira aqui o selo referente ao Objetivo do Desenvolvimento Sustentável que o trabalho atende direta ou indiretamente.

INTRODUÇÃO

O gerenciamento de irrigação é uma área importante de pesquisa em agronomia. A irrigação livre, dependente das chuvas, não é ótima pois tende a causar distribuição desigual de umidade no solo, tornando necessários métodos de monitoramento, controle e atuação sobre a irrigação (RAJKUMAR, ABINAYA, KUMAR, 2017). O planejamento inadequado da irrigação e o uso ineficiente de recursos hídricos são os dois parâmetros mais importantes restringindo a produção agricultural em diversas regiões (ROOPAEI, RAD, CHOO, 2017).

Sistemas inteligentes de irrigação dependem de um modelo baseado em memória para detectar, analisar e compensar eventos insuficientes de irrigação (GARCIA *et al.*, 2018). Os sistemas inteligentes de irrigação permitem adaptação dos cronogramas de irrigação e os tempos de funcionamento de irrigadores para atender às necessidades específicas da região sob exame.

Este trabalho trata do desenvolvimento de um robô que analisa os dados recebidos pelos sensores de umidade de solo de hortas verticais em pequena escala. Esses dados alimentam um sistema de Inteligência Artificial (IA), o qual prediz uma regra de irrigação. Os dados são fornecidos por um conjunto de sensores sendo estes: sensores higrômetros, sensor de umidade e sensor de temperatura do ambiente, DHT11 (GAY, 2014; ADAFRUIT, 2019). Estes sensores estão conectados a um *Arduino Mega*, uma placa de prototipagem eletrônica de hardware livre composta por um microcontrolador Atmel, o qual envia os dados recebidos para um computador, onde são processados e armazenados em um banco com valores separados por vírgulas (CSV). Neste sistema é possível o monitoramento da horta, registrando os dados de níveis de umidade de solo das plantas, em porcentagens e de temperatura ambiente em graus Celsius. O objetivo do sistema proposto é a irrigação inteligente, isto é, permite saber qual momento do dia é melhor para regar a horta, visando reduzir variações de umidade do solo e assim evitar o estresse hídrico da horta.

Neste contexto da proposta, objetiva-se criar um modelo que classifique os níveis de umidade do solo, obtendo as probabilidades de irrigação. Através deste modelo, pessoas desprovidas de alto capital de investimento ou recursos hídricos avançados podem automatizar suas hortas verticais e atender de forma mais adaptada cada horta, resultando em melhores resultados de produção quanto a plantio e cultivo.

O restante deste trabalho está organizado da seguinte maneira. Na seção de metodologia são descritos os métodos e materiais utilizados para a criação do sistema inteligente de irrigação e, em seguida na seção de resultados, os

resultados obtidos de predição e tomada de decisão. Na seção de conclusão são apresentadas considerações finais e possibilidades de trabalhos futuros.

METODOLOGIA

Em primeiro momento os dados de umidade e temperatura são coletados através dos sensores higrômetro e DHT11, respectivamente, usando um *Arduino Mega*, que se conecta com o computador para transmitir esses dados recebidos dos sensores, por um cabo do tipo USB. Os dados são armazenados em um banco no formato CSV, que são utilizados para as próximas etapas de pré-processamento e implementação do algoritmo que detecta o vizinho mais próximo para classificar (k-NN, *k- nearest neighbors*).

Durante o pré-processamento, são definidos valores dos dados de umidade do solo no intervalo de 0 a 100, os dados de temperatura que formam a entrada do modelo são estabelecidos a partir das leituras do ambiente. Sendo feitas quatro classificações de umidade, pré estipuladas como, árido (1-25), seco (26-50), úmido (51-75) e molhado (76-100). Com estes dados divididos, foi averiguada a possibilidade de *outliers* e não possuindo *outliers,* foi observado o comportamento de determinados dados e por inspeção visual foi determinado como estão agrupados na Imagem 1.

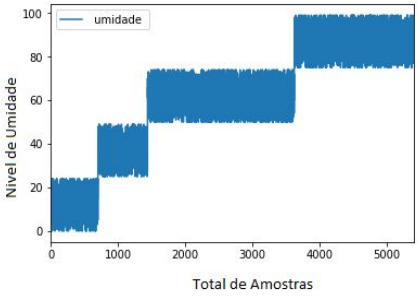

Imagem 1 - Distribuição dos dados de umidade.

As amostras catalogadas possuem dados como ID, para saber a qual mostra se refere, umidade, temperatura e data e hora que as amostras foram retiradas os dados são segmentados com X para com duas respectivas labels os dados de IDs das 4 mudas estudadas com diferentes padrões de comportamento como mostra no gráfico acima, o restando dos dados estão em Y para treino do modelo nessa parte contém os dados de umidade e temperatura.

Como os dados estão dispersos seguindo uma certa variação em padrões, a hipótese é que existe a possibilidade de usar algum algoritmo de aprendizado supervisionado usado para classificar e segmentar os dados. Para tal, foi utilizado o algoritmo k-NN para calcular as distâncias e classificar os dados.

O k-NN (algoritmo dos *k* vizinhos mais próximos) é um algoritmo simples de aprendizado de máquina que categoriza uma entrada usando seus *k* vizinhos mais próximos. O algoritmo k-NN possui várias propriedades que o diferencia de outros algoritmos de aprendizado de máquina. Primeiro, é não-paramétrico, o

que significa que, não faz nenhuma suposição sobre a distribuição de probabilidade da entrada. Isso é útil para aplicações com propriedades de entrada que são desconhecidas e, portanto, tornam o k-NN mais robusto do que os algoritmos paramétricos. O contraste é que, algoritmos paramétricos de aprendizado de máquina tendem a produzir menos erros do que os não-paramétricos, tendo em vista, as probabilidades de entrada que podem influenciar a tomada de decisão.

Segundo Ho e Peyne [2001], este teorema pode ser resumido como: *"é impossível, criar uma estratégia de otimização que seja de uso universal"*. Além disso, a única forma de uma estratégia superar outra é, se é especializada na estrutura de um problema específico em questão.

Os dados foram armazenados em *data frames* com cinco características, sendo elas, mês, dia, hora, umidade e temperatura e foi utilizado o agrupamento para descobrir qual vizinho comunica-se com os vizinhos mais próximos e passa-se a agrupar os dados em quatro grupos. Inicialmente foi aplicado aos dados um classificador k-NN com quatro componentes principais. Após o treino e teste do classificador, é feita a previsão dos dados.

Com o classificador k-NN treinado e classificando os dados de umidade, a rede de computadores opera em conjunto com o dispositivo. É utilizada uma plataforma de *Internet of Things*, a *ThingsBoard* (ELEYAN, SCOTT, 2019) a qual permite o gerenciamento de dispositivos IoT e criar telas conhecidas como dashboards para monitoramento dos sensores. A plataforma ThingsBoard foi utilizada para monitoramento em tempo real dos dados da horta e saber qual tomada de decisão o modelo está realizando em relação à plantação. Na Imagem 2 é apresentada a implementação da dashboard em funcionamento.

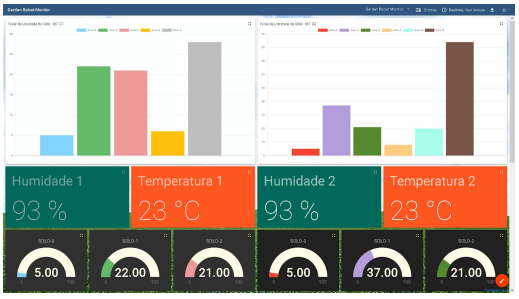

Imagem 2 - Dashboard de monitoramento da horta.

Com o protocolo de rede MQTT (*Message Queuing Telemetry Transport*), os dados são transmitidos da horta para o servidor onde fica hospedado a ThingsBoard. O MQTT é um protocolo comumente usado na indústria de coleta de dados e foi desenvolvido pela IBM (HUNKELER, TRUONG, STANFORD-CLARK, 2008) para comunicação entre mainframes. Por ser de fácil implementação é um protocolo amplamente utilizado e com variações em forma de otimização.

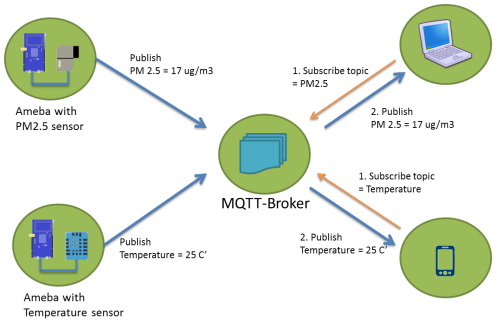

Imagem 3 - Sistema de Pub e Sub.

O MQTT opera utilizando conceito de tópicos para envio e recebimento de mensagens, assim como um conceito de topologia de rede muito semelhante a rede estrela, onde tem-se um *broker* (server), e vários dispositivos clientes IoT conectados e enviando e recebendo mensagens, sendo as mensagens nas formas *published* (enviar) e *subscribe* (receber). Os dados dos sensores são enviados usando publicadores com mqtt para um servidor da utfpr, que possui um sistema de gerenciamento de dispositivos que recebe e interpreta esses dados por meio de troca mensagens essa informação e então sobrescrevida por tópicos e para os dispositivos de visualização seja eles web aplicações ou aplicativos android.

RESULTADOS

O sistema atingiu precisão de predição de grupos de umidade de 99.68%. A base de dados criada possui um total de 5404 de amostras coletadas pelos sensores. Aplicando o agrupamento sobre o conjunto de dados de coleta foi possível identificar os grupos com seus respectivos níveis de umidade. O sistema foi capaz de encontrar e predizer a um novo sensor que ele se encaixa em um determinado grupo já classificado.

Segundo estas informações é possível saber quando é o melhor momento para a regra do plantio, porém, ao analisar os sensores com relação à onde eles foram postos de forma geográfica, fica claro que para determinados tipos de solo é factível que seja mais úmido ou mais seco para determinadas plantas, porém este dado foi apenas observado. Ainda há de se considerar que é preciso mais análises com dados distintos para novas correlações. Contudo apenas com esse simples robô é possível otimizar a quantidade de irrigação e saber quando é de fato necessário regar.

CONCLUSÃO

Levando em consideração esses aspectos, é possível analisar a aquosidade da gleba baseando-se na construção de um dispositivo IoT.

Considerando todo o cenário que engloba o contexto da proposta foi concluído que o robô pode coletar, analisar, classificar e prever o estado do solo, e baseado num modelo treinado saber quais amostras pertencem a quais grupos e com isso gerar uma irrigação autônoma analisando os parâmetros da amostra selecionada. Apesar disso, existe a necessidade de abranger novos parâmetros para comparação, com dados distintos, correlacionando com dados meteorológicos, geográficos e fluviais.

Trabalhos futuros serão concentrados em análises por séries temporais utilizando modelos como ARX, ARMAX, Box-Jenkins e *Output-Error*, assim como análises de heterocedasticidade, correlação, normalidade e distribuição. Serão consideradas também outras entradas para o modelo de tomada de decisão. Há a possibilidade de investigar-se métodos alternativos para a tomada de decisão, como árvores de decisão, regressão logística, redes neurais artificiais e modelos bayesianos.

REFERÊNCIAS

ADAFRUIT. DHT11, DHT22 and AM2302 Sensors, Datasheet, 2019.

Ho e Peyne(2001)Y.C. Ho e D.L. Peyne. Simple Explanation of the No Free Lunch Theorem of Optimization. Proceedings of the 40th IEEE Conference on Decision and Control, volume 5, páginas 4409–4414. IEEE.Citado na pág.12

GARCIA, A.M., GARCIA, I.F., POYATO, E.C. BARRIOS, P., DIAZ, J. A. Coupling irrigation scheduling with solar energy production in a smart irrigation management system. Journal of Cleaner Production, 175, pp. 670–682, 2018.

GAY, W. Experimenting with Raspberry Pi, New York: Springer, 2014.

HUNKELER, U., TRUONG, H.L., STANFORD-CLARK, A. MQTT-S — A publish/subscribe protocol for Wireless Sensor Networks. 2008 3rd International Conference on Communication Systems Software and Middleware and Workshops (COMSWARE '08), Bangalore, India, jun., 2008.

RAJKUMAR, M. N.; ABINAYA, S.; KUMAR, V. V. Intelligent irrigation system — An IOT based approach, 2017 International Conference on Innovations in Green Energy and Healthcare Technologies (IGEHT), Coimbatore, pp. 1-5, 2017.

ROOPAEI, M.; RAD, P.; CHOO, K.-K. R. Cloud of Things in Smart Agriculture: Intelligent Irrigation Monitoring by Thermal Imaging. IEEE Cloud Computing, 4(1), pp. 10-15, 2017.