

Desenvolvimento de ensembles para predição da taxa de progresso da ferrugem do cafeeiro durante seu período crítico de progresso

Thiago Toshiyuki Thamada¹
Carlos Alberto Alves Meira²

A ferrugem, provocada pelo fungo *Hemileia vastatrix* Berk. et Br., é a principal doença que atinge o cafeeiro. Pode ser encontrada em todas as regiões produtoras de café do Brasil e causa danos entre 35% e 50% na produção caso as condições climáticas favoreçam sua epidemia (ZAMBOLIM et al., 2002).

O cafeeiro possui produção bienal, alterna anos de alta e baixa carga pendente de frutos. A ferrugem ataca com maior intensidade nos anos de carga alta. A doença costuma ser controlada por meio de fungicidas, onde suas aplicações seguem um calendário fixo com início em dezembro (ZAMBOLIM et al., 2002).

Mudanças climáticas durante o outono e inverno podem atrasar o começo da epidemia e permitir que a ferrugem mantenha altos índices ao fim de seu ciclo, em agosto. Desta forma, o ciclo de epidemia é prolongando, sendo necessário rever e readaptar as medidas de controle (CHALFOUN et al., 2001). Um período crítico para o progresso da ferrugem pode ser definido de dezembro a junho.

Um posicionamento adequado das aplicações de fungicidas permite um controle eficiente da doença, racionalização de agroquímicos e preservação do meio ambiente (BETTIOL; GHINI, 2001). Nesse sentido, sistemas de alerta de doenças de plantas podem auxiliar na tomada de decisão so-

¹ Universidade Estadual de Campinas (Unicamp)

² Embrapa Informática Agropecuária

bre o controle da epidemia. Esses sistemas predizem, a partir de modelos preditivos, quando a intensidade de uma doença deverá atingir índices críticos.

Alguns modelos preditivos e sistemas de alerta para a ferrugem do cafeeiro foram desenvolvidos usando dados de todo o ano agrícola, setembro a agosto. Esses modelos predizem se, para o próximo mês, a taxa de progresso da ferrugem será igual ou maior a um determinado limiar. A taxa de progresso é a diferença entre os valores de incidência de dois meses subsequentes (GIROLAMO NETO, 2013; MEIRA et al., 2009).

É natural de se pensar em novas formas para criar modelos preditivos de desempenho preditivo superior aos existentes. O ensemble, combinação de diversos modelos, pode ser uma ferramenta capaz de criar modelos de bom desempenho. A predição inferida por um ensemble é o resultado de uma votação simples ou com peso entre os modelos que o compõe.

Esta pesquisa, ainda em execução, tem como hipótese que a criação de ensembles, a partir de dados meteorológicos e de incidência da ferrugem durante o período crítico para o progresso da doença, que determinem o aumento da taxa de progresso, proporcionará um desempenho preditivo superior aos modelos atuais.

Neste trabalho será realizada uma instância de descoberta de conhecimento em bases de dados *Knowledge Discovery in Databases* (KDD), a partir de dados meteorológicos e de incidência da doença. A Figura 1 mostra uma visão macro do processo que será utilizado para desenvolver o ensemble e obter a predição sobre o aumento da taxa de progresso da ferrugem.

A coleta dos dados que serão usados nesta pesquisa ocorreu em três cidades mineiras (Boa Esperança, Carmo de Minas e Varginha) e foi realizada pela Fundação Procafé. Os dados de incidência da ferrugem foram coletados mensalmente em lavouras com alta ou baixa carga pendente de frutos. Enquanto que os dados meteorológicos foram medidos em intervalos de trinta minutos por estações meteorológicas automáticas.

Serão usados apenas elementos meteorológicos de maior relevância para o desenvolvimento da ferrugem do cafeeiro como: temperatura, precipitação e umidade relativa do ar, visando obter modelos simples e melhor interpretação de seus resultados. A partir dessas variáveis serão derivados novos atributos considerando os períodos de incubação, de molhamento foliar e de infecção. O intuito é realizar um processo de preparação de dados distinto

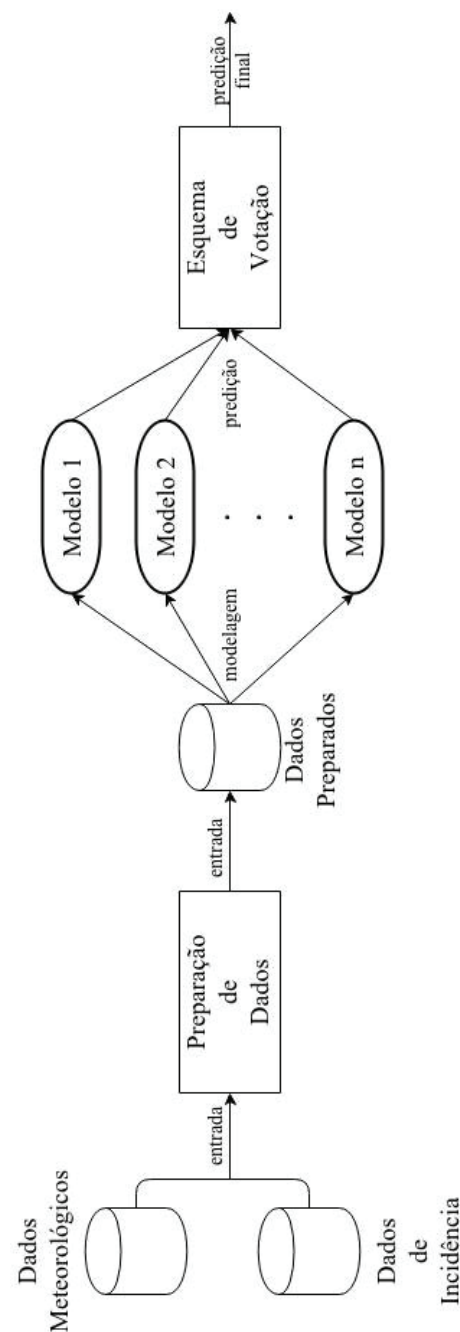


Figura 1. Esquema do processo de desenvolvimento de ensemble.

do realizado na criação dos modelos preditivos atuais, de Di Girolamo Neto (2013) e Meira et al. (2009).

Os dados climáticos deverão ser transformados de horários para mensais, para que se possa relacionar o clima com a variável a ser predita (atributo meta), que será a taxa de progresso da ferrugem na lavoura. O atributo meta será binário: caso o modelo indique aumento na taxa de progresso da ferrugem maior ou igual a um determinado limiar, a saída é 1; caso contrário, é 0.

Os processos de preparação e transformação de dados serão realizados por *scripts* na linguagem de programação Python. Os dados que formarão os conjuntos de treinamento estarão relacionadas ao período crítico para progresso da ferrugem, dezembro a junho.

Técnicas como árvore de decisão, floresta aleatória, máquina de vetor suporte (do inglês, *Support Vector Machine*) e rede neural artificial serão usadas na criação dos modelos que servirão como base para os ensembles. A predição final, a ser realizada pelos ensembles, deverá ser o resultado de uma votação entre as predições dos demais modelos. O desenvolvimento dos modelos e ensembles será realizado pelos *scripts* na linguagem de programação R.

Espera-se com essa pesquisa obter ensembles com desempenho preditivo melhor que os modelos de Di Girolamo Neto (2013) e Meira et al. (2009) durante o período crítico para o progresso da ferrugem do cafeeiro. Para isso, os modelos atuais deverão ser avaliados somente no período delimitado e seus resultados comparados com os obtidos pelos ensembles.

Posteriormente, pretende-se incorporar os ensembles em um sistema de alerta para ferrugem do cafeeiro. Assim, deseja-se que os ensembles possam auxiliar na tomada de decisão sobre o controle da doença, principalmente em anos de ocorrência da ferrugem tardia, indicando possíveis momentos oportunos para aplicação de fungicidas.

Palavras-chave: *Hemileia vastatrix*, café, modelagem de doença de planta, mineração de dados, KDD.

Referências

BETTIOL, W.; GHINI, R. Proteção de plantas em sistemas agrícolas alternativos. In: MICHEREFF, S. J.; BARROS, R. (Ed.) **Proteção de plantas na agricultura sustentável**. Recife: Universidade Federal Rural de Pernambuco, 2001. p. 1-13.

CHALFOUN, S. M.; CARVALHO, V. L. de; PEREIRA, M. C. Efeito de alterações climáticas sobre o progresso da ferrugem (*Hemileia vastatrix* Berk. & Br.) do cafeeiro (*Coffea arabica* L.). **Ciência e Agrotecnologia**, v. 25, n. 5, p. 1248-1252, set./out. 2001.

DI GIROLAMO NETO, C. **Desenvolvimento e avaliação de modelos de alerta para a ferrugem do cafeeiro**. 2013. 155 f. Dissertação (Mestrado em Engenharia Agrícola) – Faculdade de Engenharia Agrícola, Universidade Estadual de Campinas, Campinas.

MEIRA, C. A. A.; RODRIGUES, L. H. A.; MORAES, S. A. de. Modelos de alerta para o controle da ferrugem-do-cafeeiro em lavouras com alta carga pendente. **Pesquisa Agropecuária Brasileira**, v. 44, n. 3, p. 233-242, mar. 2009.

ZAMBOLIM, L.; VALE, F. X. R.; COSTA, H.; PEREIRA, A. A.; CHAVES, G. M. Epidemiologia e controle integrado da ferrugem-do-cafeeiro. In: ZAMBOLIM, L. (Ed.) **O estado da arte de tecnologias na produção de café**. Viçosa, MG: Suprema Gráfica e Editora, 2002. p. 369-449.