

Desvendando o software Fuzzy Inference
System Professional – FisPro

FUZZY
LOGIC

**Empresa Brasileira de Pesquisa Agropecuária
Embrapa Informática Agropecuária
Ministério da Agricultura, Pecuária e Abastecimento**

DOCUMENTOS 169

Desvendando o software *Fuzzy Inference System Professional* – FisPro

*Luiz Manoel Silva Cunha
João Francisco Gonçalves Antunes
Maria Fernanda Moura*

Autores

Exemplares desta publicação podem ser adquiridos na:

Embrapa Informática Agropecuária

Av. Dr. André Tosello, 209 - Cidade Universitária
Campinas, SP, Brasil
CEP. 13083-886
Fone: (19) 3211-5700
www.embrapa.br
www.embrapa.br/fale-conosco/sac

Comitê Local de Publicações
da Unidade Responsável

Presidente
Stanley Robson de Medeiros Oliveira

Secretária-Executiva
Maria Fernanda Moura

Membros
Adriana Farah Gonzalez, membro nato, Alexandre de Castro, membro indicado, Carla Cristiane Osawa, membro nato, Debora Pignatari Drucker, membro eleito, Ivan Mazoni, membro eleito, João Camargo Neto, membro indicado, João Francisco Gonçalves Antunes, membro eleito, Magda Cruciol, membro nato.

Revisão de texto
Adriana Farah Gonzalez

Normalização bibliográfica
Carla Cristiane Osawa

Projeto gráfico da coleção
Carlos Eduardo Felice Barbeiro

Editoração eletrônica
Mariana Pilatti sob supervisão de Magda Cruciol

Arte da capa
Magda Cruciol

1ª edição
Versão digital (2020)

Todos os direitos reservados.

A reprodução não autorizada desta publicação, no todo ou em parte, constitui violação dos direitos autorais (Lei nº 9.610).

Dados Internacionais de Catalogação na Publicação (CIP)

Embrapa Informática Agropecuária

Cunha, Luiz Manoel Silva.

Desvendando o software Fuzzy Inference System Professional – FisPro / Luiz Manoel Silva Cunha, João Francisco Gonçalves Antunes, Maria Fernanda Moura. - Campinas : Embrapa Informática Agropecuária, 2020.

PDF (27p.) : il. color. - (Documentos / Embrapa Informática Agropecuária, ISSN 1677-9274 ; 169).

1. Sistema de inferência fuzzy. 2. Lógica fuzzy. 3. Software FisPro. I. Antunes, João Francisco Gonçalves. II. Moura, Maria Fernanda. III. Título. IV. Embrapa Informática Agropecuária. V. Série.

CDD (21. ed.) 006.3

Autores

Luiz Manoel Silva Cunha

Estatístico, Mestre em Ciência da Computação e Matemática Computacional, Analista da Embrapa Informática Agropecuária, Campinas, SP.

João Francisco Gonçalves Antunes

Estatístico e Bacharel em Matemática aplicada e Computacional, Doutor em Engenharia Agrícola, Pesquisador da Embrapa Informática Agropecuária, Campinas, SP.

Maria Fernanda Moura

Estatística, Doutora em Ciência da Computação, Pesquisadora da Embrapa Informática Agropecuária, Campinas, SP.

Apresentação

Seres humanos são capazes de lidar com processos bastante complexos, baseados em informações imprecisas ou aproximadas. Para apoiar esses processos, os Sistemas de Inferência Fuzzy Inference Systems (FIS) têm se mostrado ferramentas bastante apropriadas.

Este documento apresenta, no formato de tutorial, conceitos sobre FIS e sua estrutura, o software Fuzzy Inference System Professional (FisPro) e um estudo de caso demonstrando a utilização de alguns recursos contidos neste software.

O FisPro mostrou ser de fácil utilização e disponibiliza recursos suficientes para o desenvolvimento de FIS. Por ser implementado em código aberto, este software traz a oportunidade de associação de outros recursos para a avaliação de desempenho dos modelos de classificação gerados.

A incorporação de métodos seleção de atributos, de novas alternativas para o estabelecimento de modelos de classificação e de regressão, e a integração do software FINGRAMS, que permite avaliação visual do comportamento das regras, tornará o FisPro uma ferramenta mais completa para o desenvolvimento de FIS.

Silvia Maria Fonseca Silveira Massruhá

Chefe-Geral

Sumário

Introdução.....	9
Sistema de Inferência <i>Fuzzy</i>	10
Software FisPro.....	10
Estudo de Caso.....	13
Considerações Finais.....	24
Referências.....	24

Introdução

O Sistema de Inferência *Fuzzy Inference System* (FIS) (Guillaume; Charnomordic, 2012a) é uma das aplicações mais conhecidas da Lógica *Fuzzy* (LF) (Zadeh, 1965), pela capacidade de incorporar os conhecimentos oriundos dos seres humanos com suas nuances, bem como expressar o comportamento do sistema desenvolvido para atender a uma demanda, de forma interpretável para os interessados. A LF é uma interface poderosa entre os ambientes numéricos e os simbólicos, especificamente por meio do conceito de variáveis linguísticas. (Zadeh, 1965).

Os FIS têm mostrado bons recursos para o mapeamento de entrada e saída no sistema, por satisfazer a propriedade de aproximação universal e, por permitirem que sejam projetados por especialistas do domínio ou por métodos dirigidos por dados. O mecanismo de inferência dos sistemas implementa o processo, no qual uma possível conclusão imprecisa seja deduzida de uma coleção de premissas imprecisas. Isto fornece um bom modelo para a representação e manipulação de um amplo espectro de informações expressadas linguisticamente (Guillaume; Charnomordic, 2011).

A possibilidade de o modelo gerado pelo FIS ser expresso utilizando regras, no formato de SE... ENTÃO...SENÃO... por meio de *Fuzzy Decision Trees* (FDT), Árvore de Decisão *Fuzzy* (Cintra et al., 2013), dá ao sistema uma maior condição de interpretabilidade, ou seja, transparência e inteligibilidade (Guillaume; Charnomordic, 2011).

A primeira abordagem para a construção de Sistemas *Fuzzy*, baseados em regras, focou-se na habilidade da LF para modelar a linguagem natural (Lima, 2015). Em seguida, apostou-se no aprendizado automático utilizando dados, por meio de métodos *Fuzzy* (Hüllermeier, 2005) e Mineração de Dados (Kruse et al., 1999). A aproximação das habilidades da LF aos métodos de aprendizado automático fez surgir uma nova visão para construção de Sistema *Fuzzy*, a *Computing with words* (CW).

A CW é uma metodologia que contrasta com o senso comum da computação, ou seja, a manipulação de números e símbolos. Ela fornece meios que podem aproximar a lacuna entre os mecanismos do cérebro humano e os processos da máquina para resolver problemas. Como resultado, ferramentas são incorporadas aos computadores para lidar com imprecisões, incertezas e verdades parciais (Herrera et al., 2009).

Neste contexto, este trabalho apresenta conceitos sobre FIS, sua estrutura e como materializá-lo utilizando o software FisPro. Um estudo de caso, envolvendo classificação de pastagens degradadas, foi desenvolvido como forma de demonstrar o uso destes conceitos e do software. O trabalho é de cunho introdutório, podendo ser expandido em trabalhos futuros, conforme o interesse. Cabe destacar que a avaliação da qualidade dos resultados extraídos do estudo de caso e, uma comparação entre o FisPro e softwares similares, não fizeram parte dos objetivos deste trabalho.

As motivações que levaram à realização deste exercício foram: a) oportunizar a avaliação e a utilização de conceitos e ferramentas para tratamento e análises de dados imprecisos ou incompletos, oriundos de experimentos de campo e fornecidos por especialistas de um determinado domínio do conhecimento; b) introduzir no escopo do projeto Mapeamento de níveis de degradação de pastagens no bioma Cerrado por meio de geotecnologias (GeoPastoCerrado)¹, o uso de diferentes recursos automáticos que possam auxiliar a definição de uma metodologia e na estruturação de

¹ Projeto aprovado no Sistema de Gestão da Pesquisa da Embrapa e disponível na plataforma ideare, sob o código SEG 10.18.03.062.00.00.

um sistema para a identificação de níveis de degradação e classificação de pastagens degradadas, de forma a fornecer respostas, facilmente interpretadas, para diferentes questionamentos oriundos do usuário final, do domínio Pastagem Degradada, inicialmente; e c) capacitar membros da equipe do projeto no uso do FisPro visando a construção de aplicações envolvendo dados imprecisos e incompletos.

Para que o leitor deste documento possa extrair maiores proveitos do FisPro, é desejável conhecimentos básicos em Lógica *Fuzzy* e Mineração de Dados (Kruse et al., 1999; Witten et al., 2011).

Este documento está organizado como segue: a) a seção 2 apresenta a arquitetura do FIS e uma descrição sucinta de seus componentes; b) na seção 3 são descritas as principais funcionalidades do software FisPro; c) na seção 4 é feita a apresentação de um estudo de caso; e d) na seção 5 são relacionadas as considerações finais.

Sistema de Inferência *Fuzzy*

Um FIS incorpora diferentes recursos encontrados na LF que foi concebida como uma extensão do modelo lógico preconizado na Lógica Booleana, na qual as sentenças são classificadas somente como verdadeiras ou falsas (1 ou 0). Na LF, conjuntos *fuzzy* são definidos por uma função de pertinência que pode assumir infinitos “graus de pertinência” ao conjunto, entre os valores 0 e 1. Na Figura 1 é apresentada a arquitetura de um FIS.



Figura 1. Componente do FIS.
Fonte: Villanueva (2020).

Dados precisos resultantes de algum sistema de medição são introduzidos no sistema e, estes, são mapeados para conjuntos *fuzzy* de entrada relevantes. O processo denominado “fuzzificação” é o responsável por transformar estes dados nestes valores em seus correspondentes nos conjuntos *fuzzy*. O mecanismo de inferência, por sua vez, opera a implicação lógica dos conjuntos *fuzzy* de entrada em uma base de regras que pode ser fornecida por especialistas e/ou extraídas de dados. O resultado alcançado na fase anterior é passado para o processo de “defuzzificação”, que produz saídas precisas. Alguns FIS podem ser expressos na forma de FDTs e “Regras de Decisão *Fuzzy*”.

Software FisPro

O FisPro (Guillaume; Charnomordic, 2012a) permite o desenvolvimento de um FIS, podendo este ser utilizado como ferramenta de modelagem de sistemas físicos ou biológicos. Este tipo de sis-

tema permite integrar dados, obtidos em experimentos de campo ou laboratoriais, com aqueles fornecidos pelos especialistas do domínio e, praticar a CW. Muitos métodos de aprendizado são do tipo “caixa-preta”, dificultando assim, a compreensão dos resultados por parte dos interessados. No FisPro, restrições são impostas aos algoritmos para tornar o raciocínio baseado em regras, facilmente interpretável, permitindo assim, obter uma melhor compreensão do funcionamento do sistema.

Este software é composto por duas partes distintas: a) uma biblioteca de funções escrita na linguagem de programação C++ que pode ser utilizada independentemente; e b) uma interface gráfica construída na linguagem Java. O FisPro é de código aberto e domínio público², e pode ser executado nos ambientes Windows, Linux, R e Docker. A Figura 2 exibe a tela inicial do software.

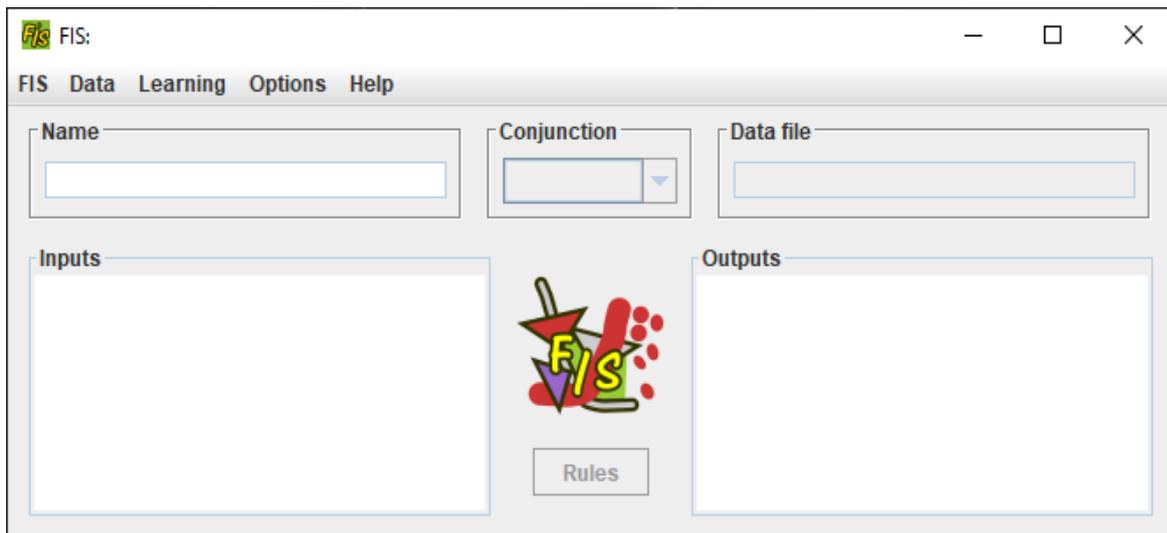


Figura 2. Conjuntos Tela inicial do FisPro.

O menu principal do FisPro é formado pelas opções FIS, Data, Learning, Options e Help (Guillaume; Charnomordic, 2018). Exceto a opção Help, as demais opções possuem submenus. Os menus e seus respectivos submenus são descritos detalhadamente em FisPro (2018) e resumidos a seguir:

FIS: este menu é formado por quatro grandes blocos. O primeiro registra opções para abrir um arquivo existente ou um novo arquivo, salvar, renomear e fechar este arquivo. No segundo bloco, temos as opções para definir as entradas e as saídas para o FIS. O terceiro bloco apresenta as opções para visualizar o resultado da inferência *fuzzy* produzida pelo sistema, que pode ser expresso como um valor exato ou no formato gráfico, em duas ou três dimensões. O último bloco deste menu fornece opções para: a) construção de FIS sem regras; b) geração de regras; e c) apresentação das conclusões para a regras. Para uso da opção de número 1, um arquivo de dados deve estar em uso. Esta opção também é empregada como referência para criar as partições *fuzzy* (Moral, 2007) e, definir as entradas e saídas para o FIS. A utilização da opção 2 exige que um FIS esteja ativo, com pelo menos uma variável de entrada. Por fim, a opção 3, para ser executada requer que um arquivo de dados esteja em uso. Uma característica importante desta opção é a possibilidade de utilizar os recursos ofertados no aprendizado do FisPro, limitado à construção de pequenos sistemas em termos do número de variáveis e de regras. É possível também desenvolver um FIS para modelar sistemas de diversas naturezas.

² Disponível em: <<https://www.fispro.org/en/>>

Data: a utilização deste menu requer que um arquivo de dados já tenha sido carregado no FisPro. Cada linha do arquivo representa uma observação, composta de n variáveis de entrada e de saída, todas elas separadas por vírgula, e a primeira linha do arquivo registra os nomes das variáveis – um formato CSV padrão. Este menu encontra-se dividido em três grandes blocos. No primeiro estão as opções para abrir e fechar arquivos de dados e extrair amostras de dados do arquivo de dados base. O segundo bloco de opções permite a elaboração de histogramas e a visualização dos dados no formato de tabela. O último bloco disponibiliza três funcionalidades importantes para analisar o desempenho dos modelos criados. A *Infer* dispara a função *Performance*, que permite a comparação entre os dados observados e as saídas inferidas para um conjunto de observações.

Em um problema de classificação, uma matriz de confusão (Witten et al., 2011) é apresentada para a avaliação do modelo. Links dá acesso às regras disparadas para diferentes conjuntos de observações contidas no arquivo de dados, acima de um limiar definido. É possível também visualizar os links existentes entre várias regras. A opção *Distance* gera uma matriz ($n \times n$) de todas as distâncias existentes entre as n observações (linhas) existentes no conjunto de dados.

Learning: o menu *Learning* disponibiliza opções para a prática do Aprendizado de Máquina Supervisionado (Monard; Baranauskas, 2003) simplificação e otimização do FIS. O submenu *Partitions* é formado pelos módulos *Generate FIS without rules* e *Hierarchical Fuzzy Partitioning-Membership Function*. *Generate* (HFP-MF) *FIS without rules* é usada como referência para a definição dos conjuntos (partições) *fuzzy*. A funcionalidade HFP-MF produz uma coleção de partições univariadas *fuzzy*, a partir de um conjunto de dados definido para treinamento do modelo. Recursos para configuração e edição de arquivo do tipo HFP, encontram-se disponíveis. É possível registrar e visualizar as coordenadas dos vértices (pontos) de todos os conjuntos *fuzzy*. Um recurso bastante interessante e disponível no FisPro é a possibilidade de obtenção dos índices (*Partition Coefficient* (PC) e *Entropy Coefficient* (PE), para a avaliação dos conjuntos *fuzzy* e a seleção dos mais adequados para o problema em estudo (Moral, 2007). O emprego deste recurso pode trazer melhorias no desempenho do modelo gerado. Para a indução de regras, além dos métodos tradicionais (Wang; Mendel, 1992) e FDT (Ichihashi et al., 1996), o FisPro fornece as opções *Fast Prototype* (FPA) e *Orthogonal least squares* (OLS) (Guillaume; Charnomordic, 2018). Em OLS é permitido definir a estrutura do FIS, caso ela não exista, selecionar e aperfeiçoar as regras. O método FPA, as regras e as suas conclusões são geradas em um único passo. A funcionalidade HFP-FIS cria um arquivo de configuração FIS combinando informações registradas no arquivo de configuração HFP e com os dados dos pontos dos vértices de cada conjunto *fuzzy*, definido. Incorporando estas e outras informações, o algoritmo FPA inicializa o processo de definição das regras. Também no menu *Learning* são fornecidas funções para a simplificação (*Simplification*) e o aperfeiçoamento (*Optimization*) de regras. A primeira, objetiva eliminar as variáveis de menor impacto nas regras que compõem o FIS.

Já a função *Optimization*, pelo método OLS, permite o aperfeiçoamento de todas as entradas do sistema, de todas as regras geradas e suas respectivas conclusões, bem como os padrões definidos para as funções de pertinências associados ao FIS.

Options: este menu permite alterar o idioma de apresentação da interface inclusive para a Língua Portuguesa e o formato de apresentação do FisPro.

Help: ativa a página de ajuda do FisPro, via *browser*.

About: fornece informações técnicas sobre o FisPro.

Estudo de Caso

Nesta seção é apresentado um estudo de caso voltado para a tarefa de classificação de pastagens degradadas, do município de Andradina, no estado de São Paulo. Um subconjunto de dados do projeto GeoDegrade³ foi utilizada na aplicação. A metodologia adotada apoiou-se no esquema apresentado na Figura 1. Modelos de classificação foram construídos explorando as abordagens FDT e Regras de Decisão *Fuzzy*. Como resultado final, um FIS para classificação de pastagens degradadas foi estabelecido. A avaliação dos resultados gerados pelo FIS não foi realizada por não fazer parte do escopo deste tutorial.

4.1 O conjunto de dados

O conjunto de dados é composto por 254 observações, divididas em dois grupos: a) 134 observações do capim *Brachiária Brizantha*; e b) 120 observações do capim *Brachiária Decumbens*. Cada observação é formada por 16 atributos preditores e um atributo classe que pode ser definida como: a) não degradada; b) em degradação; e c) degradada. Por exigência do FisPro, foram codificadas como 1, 2 e 3, respectivamente.

4.2 Seleção de atributos

Visando a construção de modelos de classificação facilmente apresentados e compreensíveis pelo usuário, eliminar atributos redundantes e irrelevantes, e obter uma representação reduzida do conjunto de dados, em termos de atributos, mas que produza os mesmos ou quase os mesmos resultados analíticos aplicou-se uma abordagem seleção de atributos (Chandrashekar; Sahin, 2014).

Os métodos empregados foram: a) CfsSubsetEval; b) GainRatioAttributeEval; c) InFoGainAttributeEval; d) OneRAttributeEval; e) ReliefFAttributeEval; e d) WrapperSubsetEval, disponíveis no software Weka⁴. Também adotamos o critério de que, o atributo que fosse selecionado por pelo menos 50% dos métodos utilizados, seria considerado um possível atributo preditor. Por ser um estudo de caso com foco no uso do FisPro, o resultado do processo de seleção de atributos não foi objeto de discussão com um especialista em classificação de pastagem. Os atributos selecionados são exibidos na Tabela 1.

Tabela 1. Atributos selecionados para o estudo de caso.

Atributo	Descrição	Tipo	Unidade de Medida
cobert_capim	Cobertura do capim	Numérico	%
cobert_pi	Cobertura de plantas invasoras	Numérico	%
alt_pasto	Altura do pasto	Numérico	cm
perc_cv	Percentual de capim verde	Numérico	%
sit_pastagem	Situação da pastagem	Texto	---

³ Disponível em: <<http://www.geodegrade.cnpem.embrapa.br>>.

⁴ Disponível em: <<https://www.cs.waikato.ac.nz/ml/weka/index.html>>.

4.3 Carga do arquivo de dados

Utilizando a opção Open, do menu Data, o arquivo de dados foi carregado no FisPro, veja Figura 3. A distribuição dos dados, para cada atributo, foi exibida por meio de histogramas ativados pela função View. Os dados também podem ser visualizados no formato de uma tabela, opção Table.



Figura 3. Especificação do arquivo de dado para carga no FisPro.

4.4 Geração de conjuntos de dados de treinamento e de teste

Carregados os dados no FisPro, os conjuntos de dados para treinamento e teste do modelo foram definidos, seguindo as opções *Data* → *Generate samples* → *Generate learning-test pairs*. A Figura 4 apresenta os parâmetros de configuração para esta funcionalidade. Um único par de arquivos (treinamento-teste) foi definido (*number of sample pair* = 0) para este estudo. Para obter um número maior de pares de arquivo, deve-se alterar o valor do parâmetro. Os conjuntos de dados são gerados de forma aleatória, a cada chamada. Para isso, o parâmetro *Seed* deve registrar o valor 0. A adoção dessa prática é desejável para evitar que o treinamento e o teste do modelo utilizem os mesmos dados. Caso não ocorra a divisão do conjunto de dados, como citado, os resultados gerados poderão não fornecer os conhecimentos que estão sendo buscados.

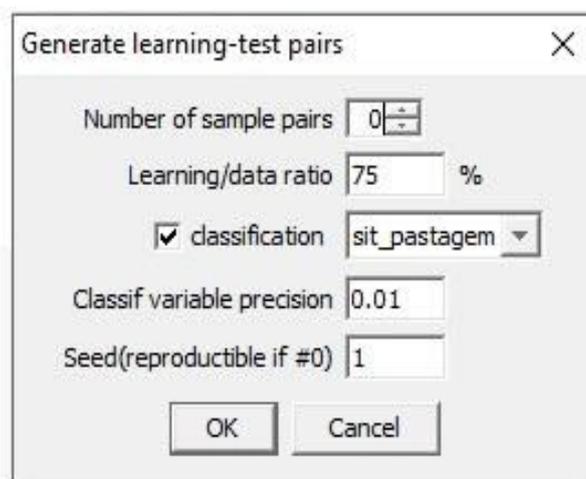


Figura 4. Parâmetros de configuração na geração dos conjuntos de dados de treinamento e teste.

A utilização de vários pares de conjuntos de dados (treinamento e teste) traz a possibilidade de obtenção de modelos para diferentes cenários. A opção Validação Cruzada não é disponibilizada neste momento, mas na fase de Otimização do sistema, conforme comentários descritos no item 4.6.2. Por *default*, o arquivo de treinamento registra 75% do número total de observações, mas este valor pode ser alterado. Ativada a opção *classification*, indica-se qual atributo é utilizado como a classe. As quantidades proporcionais de observações, em cada classe serão mantidas, nos pares

de arquivos gerados. O parâmetro *Seed* = 0 indica que novas amostras de observações serão geradas, de forma aleatória, a cada chamada desta função.

4.5 Definição dos conjuntos *fuzzy*

Seguindo a sequência de opções *Learning* → *Partitions* → *Generate FIS with*, a partir do menu principal, o processo de definição dos conjuntos *fuzzy* se inicia, como ilustrado na Figura 5.

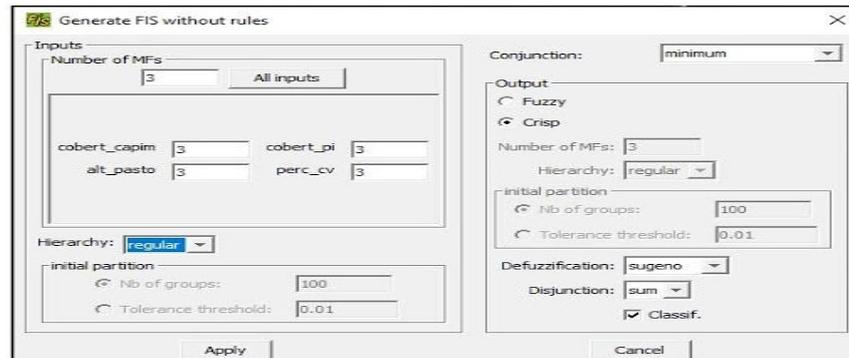


Figura 5. Definições da estrutura do FIS.

Foram definidos três conjuntos para cada variável de entrada do sistema. O tipo de *Hierarchy* adotada para a formação deles, foi a *regular*, conforme Lima (2015). O operador *minumum* (\wedge , lógico para avaliar a sentença), parâmetro *Conjunction*, foi estabelecido para conectar as sentenças a serem investigadas no FIS, isto é, $f_{A \cap B}(x) = 1$, se $x \in A$ e $x \in B$. A saída do sistema será um número *Crisp*. O método “sugeno” (Kaur; Kaur, 2012) foi o indicado para o processo de *Defuzzification*. A operação *sum* será aplicada ao processo de *Disjunction*. A opção *Classif* assinalada significa que se trata de um problema de classificação supervisionada.

Estabelecida a estrutura para o FIS, isto é, com os conjuntos *fuzzy* definidos, conforme ilustrado na Figura 6, clica-se sobre o nome de cada variável de entrada do sistema, e uma estrutura básica para o conjunto é exibida. A estrutura foi definida tomando-se como base as informações apresentadas na Figura 5. As Figuras 7 e 8 exibem os conjuntos *fuzzy* criados. A partir deste momento, a estrutura passou a se chamar *Classif_Pastagens*.



Figura 6. Estrutura para o FIS.

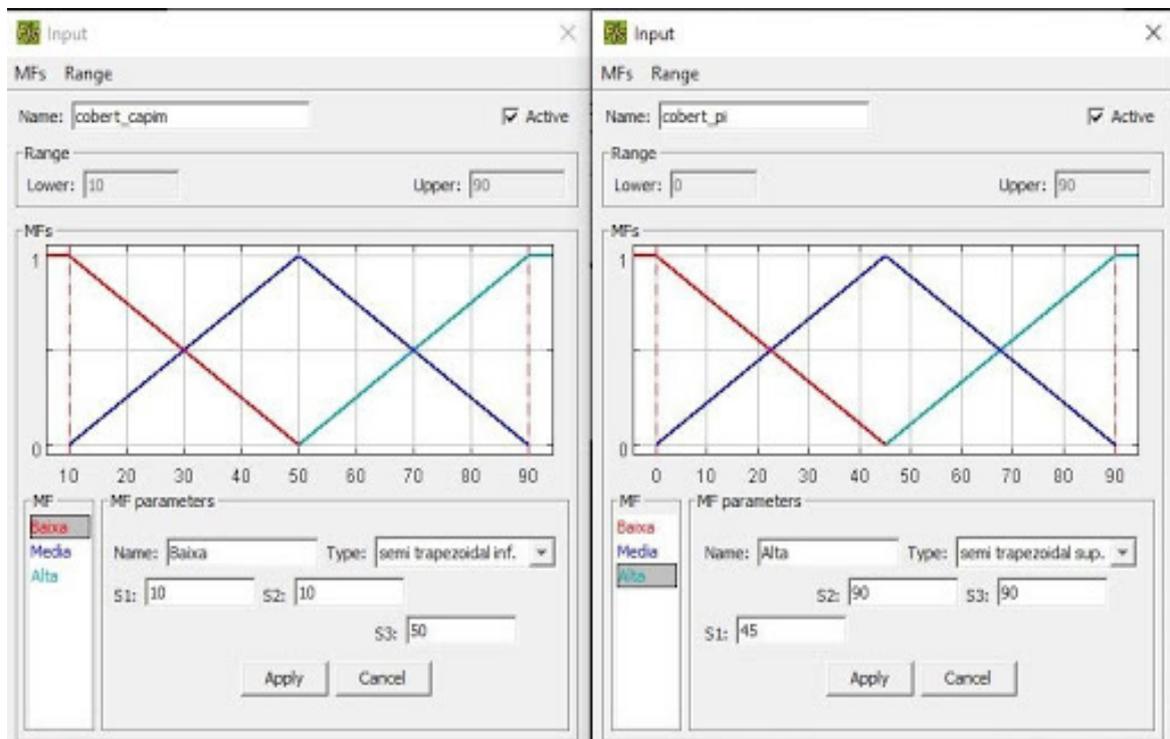


Figura 7. Conjuntos fuzzy: atributos cobert_capim e cobert_pi.

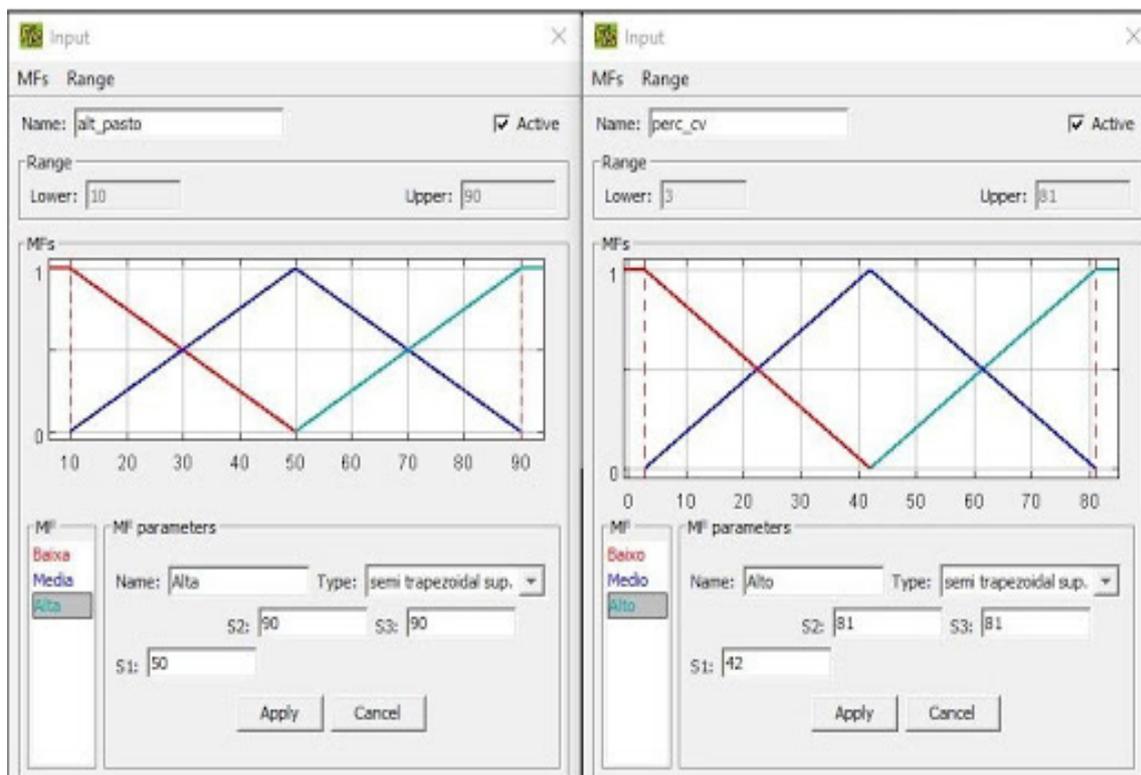


Figura 8. Conjunto fuzzy: atributos alt_pasto e perc_cv.

Acionada a opção *Range*, os valores inicial e final assumidos por cada um dos atributos são exibidos. Os conjuntos foram rotulados utilizando as variáveis linguísticas Baixo, Média/Médio e Alto, formados por uma função de pertinência semi-trapezoidal, nas extremidades e, por funções triangulares na parte central. Na Tabela 2 estão registrados os rótulos e os valores atribuídos para seus pontos inicial, central e final (S1, S2, S3) que definem as funções de pertinência. Estes valores foram extraídos dos universos de discurso definidos para cada variável, de forma automática, mas que podem ser alterados conforme a necessidade do interessado. Os formatos das funções de pertinência também podem ser modificados.

Tabela 2. Rótulos e valores atribuídos aos pontos das funções de pertinências.

Atributo: <u>cobert capim</u>			
Rótulo	S1	S2	S3
Baixo	10	10	50
<u>Media</u>	10	50	90
Alta	50	90	90
Atributo: <u>cobert pi</u>			
Rótulo	S1	S2	S3
Baixo	0	0	45
<u>Media</u>	0	45	90
Alta	45	90	90
Atributo: <u>alt pasto</u>			
Rótulo	S1	S2	S3
Baixo	10	10	50
<u>Media</u>	10	50	90
Alta	50	90	90
Atributo: <u>perc cv</u>			
Rótulo	S1	S2	S3
Baixo	3	3	42
<u>Media</u>	3	42	81
Alta	42	81	81

4.6 Estabelecendo os Modelos de Classificação

Dois modelos foram desenvolvidos, um por meio de Regras de Decisão Fuzzy e outro utilizando FDT com as opções para estabelecimento dos modelos apresentadas a seguir.

4.6.1 Modelo baseado em Regras de Decisão Fuzzy

Selecionando a opção *Wang & Mendel*, seguindo o caminho *Learning* → *Rule induction* → *Wang & Mendel*, a janela exibida na Figura 9 é exibida e, clicando no botão *Rules*, as regras foram geradas, veja Figura 10.

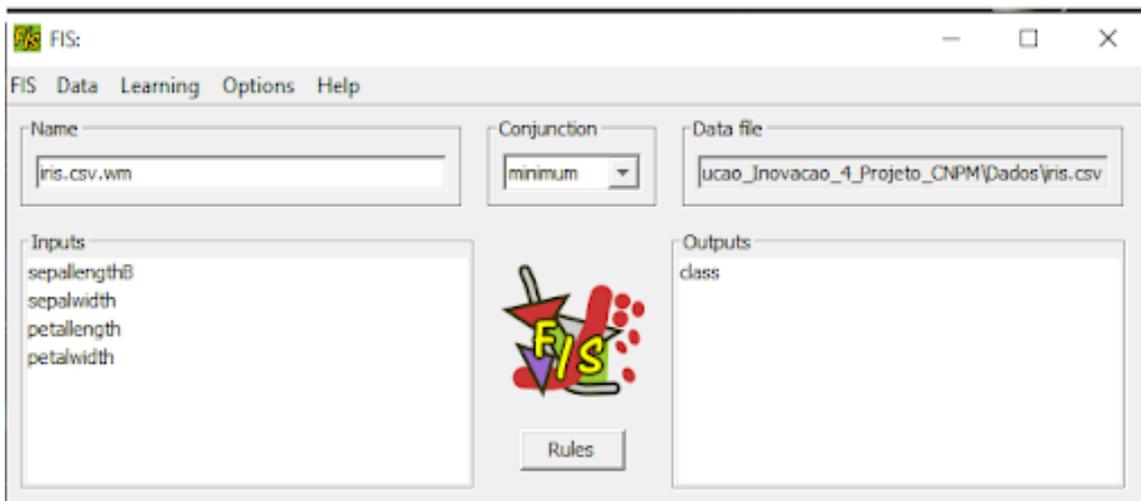


Figura 9. Geração de modelo baseado em Regras de Decisão.

Rule	Active	IF cobert_capim	AND cobert_pi	AND alt_pasto	AND perc_cv	THEN sit_pastag...
1	<input checked="" type="checkbox"/>	Alta	Media	Media	Alto	1
2	<input checked="" type="checkbox"/>	Alta	Media	Media	Baixo	1
3	<input checked="" type="checkbox"/>	Media	Media	Media	Baixo	1
4	<input checked="" type="checkbox"/>	Media	Media	Baixa	Baixo	1
5	<input checked="" type="checkbox"/>	Alta	Baixa	Media	Baixo	2
6	<input checked="" type="checkbox"/>	Media	Media	Alta	Baixo	2
7	<input checked="" type="checkbox"/>	Media	Baixa	Baixa	Baixo	3
8	<input checked="" type="checkbox"/>	Alta	Media	Baixa	Alto	3
9	<input checked="" type="checkbox"/>	Alta	Media	Baixa	Medio	1
10	<input checked="" type="checkbox"/>	Media	Media	Baixa	Medio	2
11	<input checked="" type="checkbox"/>	Media	Baixa	Baixa	Medio	3
12	<input checked="" type="checkbox"/>	Alta	Baixa	Baixa	Medio	3
13	<input checked="" type="checkbox"/>	Media	Baixa	Media	Baixo	1
14	<input checked="" type="checkbox"/>	Alta	Baixa	Alta	Medio	1
15	<input checked="" type="checkbox"/>	Media	Baixa	Media	Medio	2
16	<input checked="" type="checkbox"/>	Media	Media	Media	Medio	3
17	<input checked="" type="checkbox"/>	Alta	Baixa	Alta	Alto	1
18	<input checked="" type="checkbox"/>	Alta	Baixa	Alta	Baixo	1
19	<input checked="" type="checkbox"/>	Media	Baixa	Baixa	Alto	3
20	<input checked="" type="checkbox"/>	Alta	Baixa	Baixa	Alto	3
21	<input checked="" type="checkbox"/>	Alta	Baixa	Media	Alto	1
22	<input checked="" type="checkbox"/>	Baixa	Alta	Baixa	Baixo	3
23	<input checked="" type="checkbox"/>	Baixa	Alta	Media	Baixo	3
24	<input checked="" type="checkbox"/>	Alta	Baixa	Media	Medio	2

Figura 10. Regras de Decisão Fuzzy.

A utilização da abordagem Wang&Mendel, no FisPro, requer a definição prévia dos conjuntos *fuzzy*. O ponto desfavorável desta abordagem é que ela não é muito adequada para saídas *crisps*, quando estas representam valores numéricos contínuos. Quando esta abordagem é utilizada, a saída observada registra até 15 valores distintos. Mesmo assim, esta abordagem é adequada a problemas de classificação. O modelo de inferência de Takagi e Sugeno (1985) não é o mais indicado para classificação, e sim para regressão.

Conforme apresentado na Figura 10, ao todo, 24 regras foram geradas com suas respectivas classes no modo ativo, ou seja, são consideradas no processamento. Acionando a opção *Rules*, novas regras podem ser inseridas a este conjunto. Um peso pode ser atribuído a cada uma das regras, como forma de diferenciar o seu grau de importância, usando a opção *Display*. Por meio do menu FIS, o modelo gerado foi salvo como “Classif_Pastagens.wm”.

O FisPro também oferece o método OLS (*Ordinary Least Squares*) para geração de Regras de Decisão *Fuzzy*. É um processo mais robusto que o Wang&Mendel. Em OLS as regras são estabelecidas e as mais importantes são selecionadas pelo critério dos mínimos quadrados, por regressão linear e ortogonalização Gram-Schmidt. Na sequência, as regras são aperfeiçoadas e as conclusões definidas. (Guillaume; Charnomordic, 2018). Por exigir uma compreensão maior de todo este processo, esta opção não foi exercitada neste estudo.

O emprego de um desses métodos de geração de regras pode ser útil para a descoberta de novas regras, visando à formatação de uma base de regras para um Sistema Especialista.

4.6.2 Modelo baseado em Arvore de Decisão Fuzzy

A sequência de opções *Learning* → *Rule induction* → *Tree* → *Generate* leva à construção de uma FDT. No FisPro, o algoritmo responsável por induzir as árvores de decisão é o ID3 (Ichihashi et al., 1996), que exige que a estrutura do FIS esteja carregada no software.

Nesta etapa, os arquivos de treinamento e de teste foram utilizados na construção do modelo. Na Figura 11 são apresentados os parâmetros de configuração do ID3.



Figura 11. Janela de parametrização algoritmo ID3.

O único parâmetro que teve seu valor padrão alterado foi o *Validation*, que registra o nome do arquivo para teste do modelo. O emprego desta estratégia evita que o mesmo conjunto de dados utilizado no treinamento seja, também, utilizado para teste do modelo. Acionando o botão *Apply*, duas FDTs (modelos) foram geradas. A primeira contém todas as combinações de atributos possíveis para o estabelecimento da classificação e, a segunda, registra uma FDT ajustada (podada), ou seja, contendo uma estrutura mais enxuta para a classificação, Figura 12.

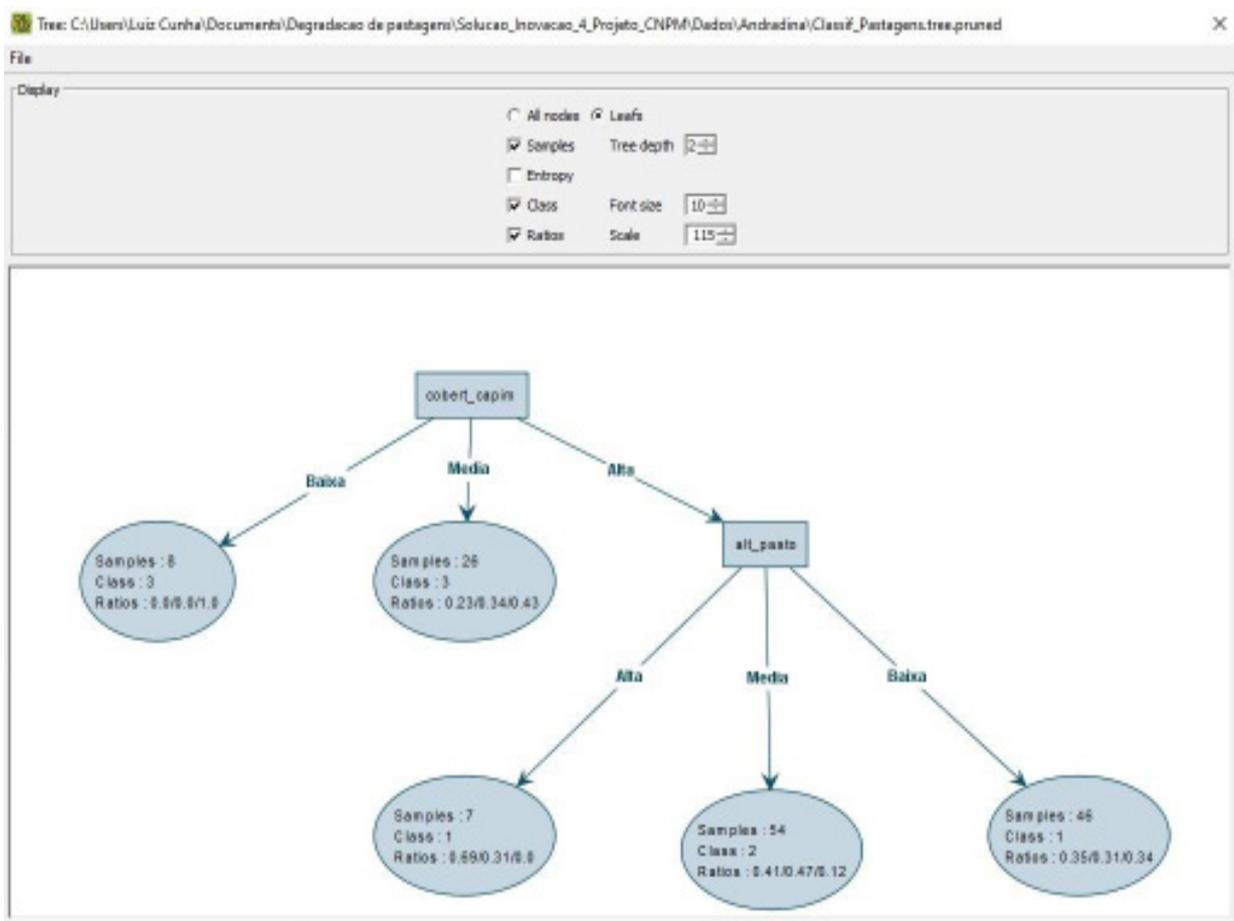


Figura 12. Árvore de Decisão podada.

Como podemos observar na Figura 11, a opção *Cross-Validation* não é disponibilizada neste momento, que é empregada quando a funcionalidade otimização do sistema (*Optimization Menu*) (Guilherme; Charnomordic, 2012b, 2018) é ativada no FisPro. No FisPro, a tarefa de otimização se dá sobre os parâmetros das funções de pertinências e conclusões das regras. O procedimento de otimização realiza a validação cruzada utilizando *10-fold*.

Dos quatro atributos preditores selecionados no item 4.2, dois deles (*cobert_capim* e *alt_pasto*) foram escolhidos para compor a FDT com cinco ramos. Ao final de cada ramo, o nó-folha registra informações referentes à quantidade de observações (*Sample*) classificadas no nó, ao valor da classe (*Class*) e à proporção de acerto (*Ratio*) para cada uma das classes (três). A informação de *Ratio* é um dos diferenciais da FDT, em relação à Árvore de Decisão (DT) tradicional (Witten et al., 2011), por apresentar os valores registrados em *Ratio* para as três classes. O valor máximo para este parâmetro é um.

Examinando os valores atribuídos ao parâmetro Ratio, (Classe = 1), o caminho cobert_capim = Alta e alt_pasto = Alto (0,69), é o que apresenta a maior proporção de acerto, logo dizemos que esta regra é a mais assertiva. As condições cobert_capim = Alta e alt_pasto = Baixa, também levam a Classe 1, mas com grau de acerto de 0,35.

Para salvar os modelos gerados, ou seja, as FDTs com e sem poda, deve-se ativar o menu FIS e, em seguida, clicar em Save as fornecendo o nome do arquivo, conforme destacado nas Figuras 13 e 14.

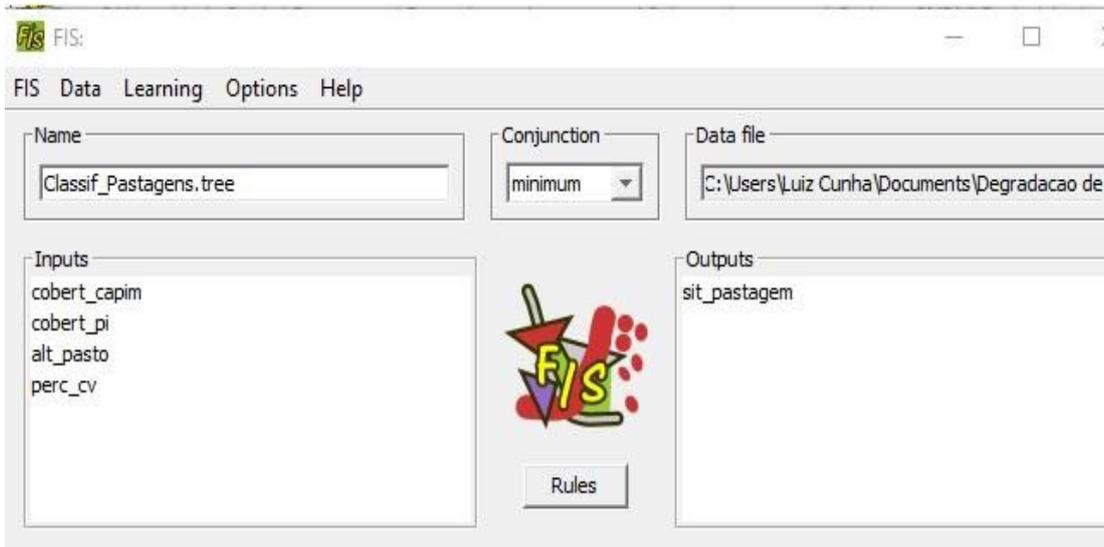


Figura 13. Janela para salvar o modelo gerado sem poda.

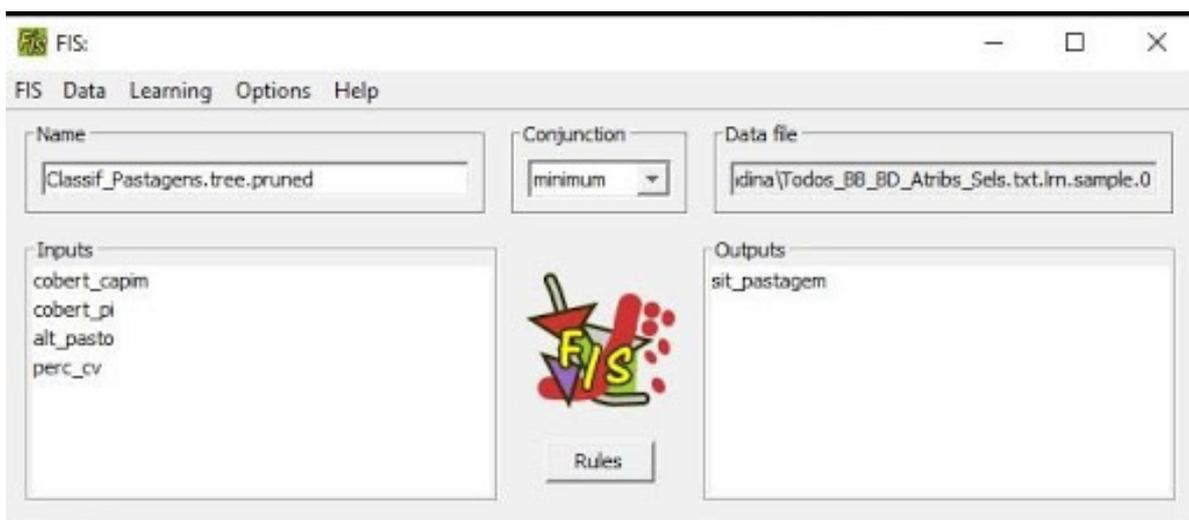


Figura 14. Janela para salvar o modelo gerado com poda.

4.7 Inferência

Esta funcionalidade é responsável por apresentar o desempenho dos modelos gerados, via *Árvore de Decisão Fuzzy*. Em linhas gerais, as saídas observadas são comparadas com as saídas inferidas, para certa amostra de dados. Para ativar esta funcionalidade seguiu-se o caminho *Data* → *Infer*, que levou aos resultados mostrados na Figura 15. A opção *Links between rules and items*, uma vez selecionada, indica a construção de um arquivo contendo os links existentes. Acionando o botão *Compute* um arquivo é gerado.

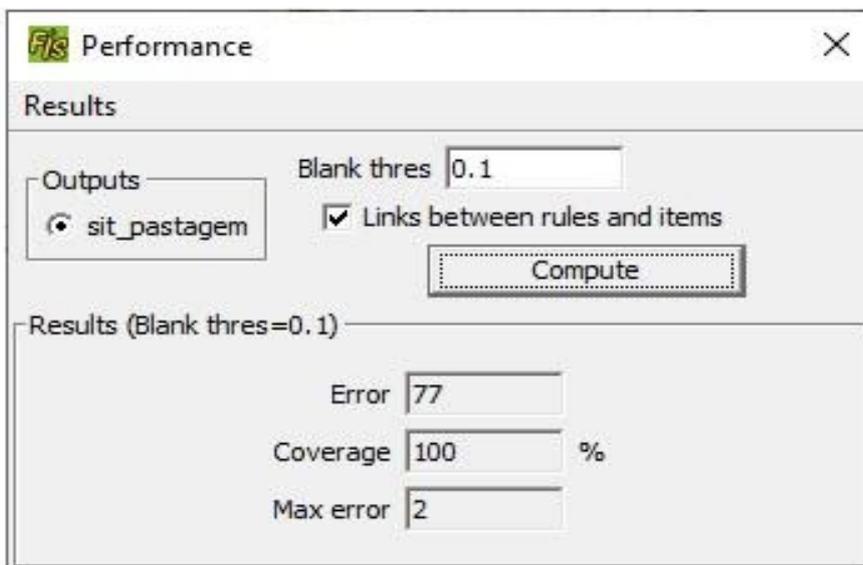


Figura 15. Apresentação de desempenho do modelo.

Em se tratando de uma tarefa de classificação, onde a última coluna das observações registra a classe, três índices complementares (Guillaume; Charnomordic, 2018) são fornecidos para avaliar o desempenho do modelo, Figura 15:

- *Error*, conhecido também como *Performance Index*, registra a quantidade de observações que tiveram as classes inferidas diferentes das classes observadas.
- *Coverage* informa a quantidade de observações coberta pelo modelo.
- *Max error*, é calculado como o valor máximo absoluto, em todo o conjunto de observações, entre a saída observada, y_i para o i -ésimo observação, e a inferida pelo sistema.

O FIS analisou todas as observações existentes no arquivo de dados de teste, encontrando 77 divergências entre as classes observadas e previstas. Acionando *Results* → *View* → *Classif.*, a Matriz de Confusão (MC) (Witten et al., 2011) do modelo é apresentada, Figura 16.

		class 1	class 2	class 3
inferred	class 1	5	0	1
	class 2	51	55	12
	class 3	6	7	40
not classified		0	0	0

Figura 16. Matriz de Confusão.

Acionando *Result* → *Save*, o arquivo contendo o resultado do desempenho do modelo é registrado.

Além dos três índices apresentados para a avaliação de modelo, as medidas Precisão, Revocação (Sensibilidade), Medida-F, Acurácia e Índice Kappa (Matos et al., 2009; Pacheco, 2018) podem ser definidas, com base nas informações contidas na Matriz de Confusão. Estas novas informações fazem com que os modelos possam ser avaliados sob outros aspectos, fornecendo melhores condições para aceitação ou rejeição do(s) modelos.

4.8 Links

A opção *Links*, no menu *Data*, permite visualizar as ligações entre as regras e as observações cobertas e, também, as conexões entre as várias regras. Esta é uma opção importante para a identificação de discordâncias entre os resultados observados e os inferidos, utilizando as regras estabelecidas. Os conteúdos destes arquivos permitem também avaliar o modelo. Quatro tipos de arquivos são gerados:

- *rules.items*, registra tantas linhas quanto a quantidade de regras no sistema, mais duas linhas.
 - › 1ª. linha, informa o número de regras geradas pelo modelo.
 - › 2ª. linha, registra o número máximo de observações que ativaram a regra.
 - › Da 3ª. linha em diante, as seguintes informações são descritas: número da regra (começando de 1, quando as regras não estão ordenadas pelo peso), o peso da regra, o número de observações que ativam a regra e os números das observações.
- *items.rules*, possui tantas linhas quanto a quantidade de linhas existentes no arquivo de dados. Cada linha é formada pelo número da observação, iniciando em 1, seguido pelos números das regras que são ativadas.
- *rules.links*, a notação de ligação entre regras é útil para verificar a consistência da base de regras. Se duas regras estão fortemente interligadas e apresentando diferentes conclusões, então há inconsistência. Esta pode ser oriunda de uma faixa de valores de cobertura da regra, mal especificada ou corresponder a uma exceção nos dados da observação.
- *rules.sorted*, se a opção *sort* está selecionada, regras são ordenadas pelo peso calculado para a regra. Isto representa sua influência no arquivo de amostras.

Considerações Finais

Este trabalho apresentou conceitos envolvidos na construção de um Sistema de Inferência *Fuzzy*, a arquitetura do sistema e um estudo de caso de aplicação de como desenvolver modelos de classificação de objetos, utilizando Regras de Decisão *Fuzzy* e FDT, utilizando recursos existentes no FisPro.

O FisPro é domínio público, de fácil manuseio e pode ser executado nos ambientes Windows, Linux, R e Docker, o que permite a inúmeros interessados obterem e utilizarem esta ferramenta. O desenvolvimento do estudo de caso, com dados reais, foi importante para avaliar o funcionamento de um conjunto de recursos oferecidos pelo software estudado e, também, por apresentar um tutorial de passo a passo orientando o desenvolvimento de um FIS.

A funcionalidade Validação Cruzada, que é praticamente um padrão em Aprendizado de Máquina, especialmente quando o sistema age sobre conjuntos pequenos de exemplos, é explorada durante a execução do procedimento de otimização do FIS.

Como apontado nos itens 4.7 e 4.8, os modelos de classificação gerados pelo FisPro podem ser avaliados empregando-se as medidas de desempenho visualizada na Figura 15, como também, utilizando medidas derivadas de informações contidas na Matriz de Confusão. O uso combinado destas medidas dá mais sustentação para a aceitação ou rejeição dos modelos. A utilização das funcionalidades existentes no FisPro, via linha de comando, é vista como uma facilidade para automatização do processamento.

O FisPro, por ser de código aberto, permite a implementação de novas funcionalidades. A incorporação de métodos de seleção de atributos evitará a utilização de ferramentas desassociadas do software FisPro. O acoplamento do pacote FRBS (Fuzzy Rules- Based Systems), para geração de Regras de Decisão *Fuzzy*, aumentará as opções para geração de FIS baseado em regras. Para a avaliação visual do comportamento de regras, encontra-se disponível o software FINGRAMS. A integração deste ao FisPro trará maiores capacidades para avaliação do FIS, e descoberta de novos conhecimentos.

Como trabalhos futuros, aprofundaremos os estudos iniciados e a exploração de outras funcionalidades contidas no FisPro, incluindo-se a avaliação da qualidade dos resultados.

Referências

CHANDRASHEKAR, G.; SAHIN, F. A survey on feature selection methods. **Computers & Electrical Engineering**, v. 40, n. 1, p. 16–28, Jan. 2014. DOI: 10.1016/j.compeleceng.2013.11.024.

CINTRA, M. E.; MONARD, M. C.; CAMARGO, H. A. A fuzzy decision tree algorithm based on C4.5. **Mathware & Soft Computing Magazine**, v. 20, n. 1, p. 56–62, 2013.

FISPRO: an open source portable software for fuzzy inference systems: version 3.6. 2018. 56 p. Disponível em: <<https://www.fispro.org/download/documentation/fispro36inline.pdf>>. Acesso em: 5 out. 2020.

GUILAUME, S.; CHARNO MORDIC, B. Learning interpretable fuzzy inference systems

with FisPro. **Information Sciences**, v. 181, n. 20, p. 4409–4427, Oct. 2011. DOI: 10.1016/j.ins.2011.03.025.

GUILAUME, S.; CHARNOMORDIC, B. Fuzzy inference systems: an integrated modeling environment for collaboration between expert knowledge and data using FisPro. **Expert Systems with Applications**, v. 39, n. 10, p. 8744–8755, Aug. 2012a. DOI: 10.1016/j.eswa.2012.01.206.

GUILAUME, S.; CHARNOMORDIC, B. Parameter optimization of a Fuzzy Inference Systems using the FisPro open source software. In: IEEE INTERNATIONAL CONFERENCE ON FUZZY SYSTEMS, Brisbane, 2012. **Proceedings...** Piscataway: IEEE, 2012b. 8 p. DOI: 10.1109/FUZZ-IEEE.2012.6250821.

GUILAUME, S.; CHARNOMORDIC, B. **FisPro**: an open source portable software for fuzzy inference systems. 2018. 56 p. Disponível em: <<https://www.fispro.org/download/documentation/fispro36inline.pdf>>. Acesso em: 3 set. 2020.

HERRERA, F.; ALONSO, S.; CHICLANA, F. HERRERA-VIDEAMA, E. Computing with words in decision making: foundations, trends and prospects. **Fuzzy Optimization and Decision Making**, v. 8, n. 4, p. 337–364, Dec. 2009. DOI: 10.1007/s10700-009-9065-2.

HÜLLERMEIER, E. Fuzzy methods in machine learning and data mining: status and prospects. **Fuzzy Sets and Systems**, v. 156, n. 3, p. 387–406, Dec. 2005. DOI: 10.1016/j.fss.2005.05.036. DOI: 10.1016/j.fss.2005.05.036.

ICHIHASHI, H.; SHIRAI, T.; NAGASAKA, K.; MIYOSHI, T. Neuro-Fuzzy ID3: a method of inducing fuzzy decision tree with linear programming for maximizing entropy and an algebraic method for incremental learning. **Fuzzy Sets and Systems**, v. 81, n. 1, p. 157–167, July 1996. DOI: 10.1016/0165-0114(95)00247-2.

KAUR, A.; KAUR, A. Comparison of Mamdani-type and Sugeno-type fuzzy inference systems for air conditioning system. **International Journal of Soft Computing and Engineering (IJSCE)**, v. 2, n. 2, p. 323–325, May 2012.

KRUSE, R.; NAUCK, D.; BORGELT, C. **Data mining with fuzzy methods**: status and perspectives. 1999. Disponível em: <<<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.97.5401&rep=rep1&type=pdf>>. Acesso em: 2 set. 2020>.

LIMA, H. P. de. **Uma abordagem para construção de sistemas fuzzy baseados em regras integrando conhecimento de especialistas e extraído de dados**. 2015. 137 f. Dissertação (Mestrado em Ciência da Computação) - Centro de Ciências Exatas e de Tecnologia - Universidade Federal de São Carlos, São Carlos.

MATOS, P. F.; LOMBARDI, L. de O.; CIFERRI, R. R.; PARDO, T. A. S.; CIFERRI, C. D. de A.; VIEIRA, M. T. P. **Relatório técnico “Métricas de Avaliação”**: Projeto “Um Ambiente para Análise de Dados da Doença Anemia Falciforme”. São Carlos, SP: UFSCar, 2009. 15 p.

MONARD, M. C.; BARANAUSKAS, J. A. **Conceitos de aprendizado de máquina**. In: REZENDE, S. O. (Coord.). **Sistemas inteligentes: fundamento e aplicações**. Barueri: Manole, 2003. p. 89–114.

MORAL, J. M. A. **Interpretable fuzzy systems modeling with cooperation between expert and induced knowledge**. 2007. 244 p. Tese (Doutorado) - Escuela Técnica Superior de Ingenieros de Telecomunicación - Universidad Politécnica de Madrid, Madrid.

PACHECO, A. **Medida de desempenho de classificadores - parte 1**. 2018. Disponível em: <<http://computacaointeligente.com.br/artigos/medidas-classificadores-1/>>. Acesso em: 2 set. 2020.

TAKAGI, T.; SUGENO, M. Fuzzy identification of systems and its applications to modeling and control. **IEEE Transactions on Systems, Man, and Cybernetics**, v. 15, n. 1, p. 116–132, 1985. DOI: 10.1109/TSMC.1985.6313399.

VILLANUEVA, J. M. **Sistema de inferência fuzzy**. Disponível em: <<https://docplayer.com.br/67741696-Sistema-de-inferencia-fuzzy-prof-juan-mauricio-villanueva.html>>. Acesso em: 10 set. 2020

WANG, L. X.; MENDEL, J. M. Generating fuzzy rules by learning from examples. **IEEE Transactions on Systems, Man, and Cybernetics**, v. 22, n. 6, p. 1414–1427, Nov. 1992.

WITTEN, I. H.; FRANK, E.; HALL, M. A. **Data mining: practical machine learning tools and techniques**. 3. ed. Amsterdam: Elsevier, 2011. 629 p.

ZADEH, L. A. Fuzzy sets. **Information and Control**, v. 8, n. 3, p. 338–353, June 1965. DOI: [https://doi.org/10.1016/S0019-9958\(65\)90241-X](https://doi.org/10.1016/S0019-9958(65)90241-X).



Informática Agropecuária