

1. INTRODUÇÃO

O conceito de casa inteligente refere-se a um ambiente residencial que, por seus sensores e atuadores conectados, é capaz de fornecer suporte inteligente e contextualizado ao usuário. Segundo FEKI, M. 2013 *et al.*, com a expectativa de que 50 a 100 bilhões de dispositivos *IoT* estejam conectadas à Internet até 2020, ocasionando uma mudança de paradigma na qual os objetos do cotidiano se tornam interconectados e inteligentes. A alta interoperabilidade de dispositivos pode levar a comportamentos indesejáveis do sistema. Indesejável aqui significa que o comportamento exibido viola os requisitos pré-estabelecidos do sistema, esse comportamento indesejável é conhecido como *Interação de Características (IC)*.

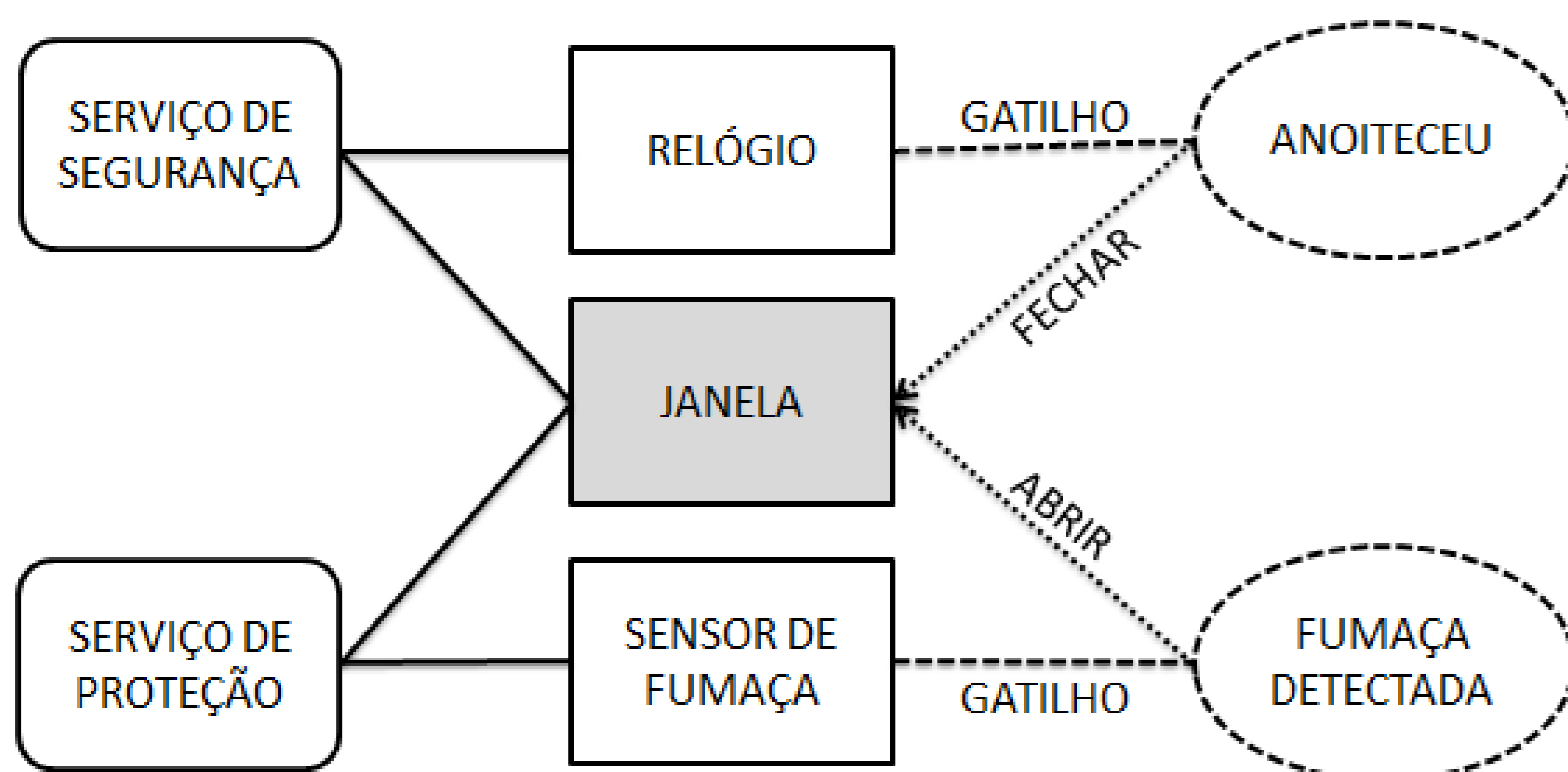


Figure 1: cenário no qual as configurações de um ambiente podem Smart-Home resultar em uma IC

2. METODOLOGIA

Deteção de IC, através de técnicas de classificação em aprendizado de máquina sobre dados simulados de um mês de atividades em um ambiente de Casa Inteligente com dispositivos interconectados que podem ou não gerar esses tipo de interações.

3. EXPERIMENTOS e RESULTADOS

Os experimentos basearam-se na seguinte composição:

- Classificadores usados: (i) Análise Discriminante Linear (LDA), (ii) Máquina de vetor de suporte (SVM), (iii) Árvore de Decisão, (iv) Floresta Randômica (Random Forest) e (v) Regressão Logística;
- A base de dados usada tem um total de 44671 observações e um total 18 variáveis, sendo 2 do tipo numérica e 16 do tipo lógica.;
- A variável resposta foi criada a partir da anotação de regras que expressam relacionamentos frequentes entre as variáveis que indicam a ocorrência; (*TRUE*), ou não (*FALSE*), de IC entre eventos do ambiente inteligente.
- A validação foi feita usando *Holdout* com 70% treinamento e 30% teste.

Table 1: Resultados das métricas de desempenho dos algoritmos utilizados

Algoritmos	Acurácia	Precisão	Recall	F1 Score	Kappa
LDA	0.9316	0.9975	0.8023	0.8893	0.8407
Reg. Logística	0.9585	0.9616	0.9155	0.9379	0.9068
Árvore de Decisão	0.9875	1.0000	0.9634	0.9818	0.9719
SVM	0.9927	1.0000	0.9787	0.9892	0.9837
Random Forest	0.9988	0.9997	0.9967	0.9982	0.9973

- O classificador *Random Forest* foi o algoritmo que teve melhor acurácia (maior taxa de acertos para a base de dados de validação);
- Todos os classificadores apresentaram coeficiente de concordância *Kappa* no intervalo entre 0.81 e 1, indicando que os dados classificados tem alta coesão;
- Todos os algoritmos apresentaram resultados de precisão altos, maiores que 96%, o que mostra que são adequados para classificar os verdadeiros positivos (ocorrência de IC);
- Para o *recall* o algoritmo LDA (80.23%) foi o único com valor abaixo dos 90% mostrando que entre os algoritmos usados ele é mais propício a gerar falsos positivos. Este menor valor de *recall* gerou um F1-score abaixo dos 90% para o LDA (88.93%).

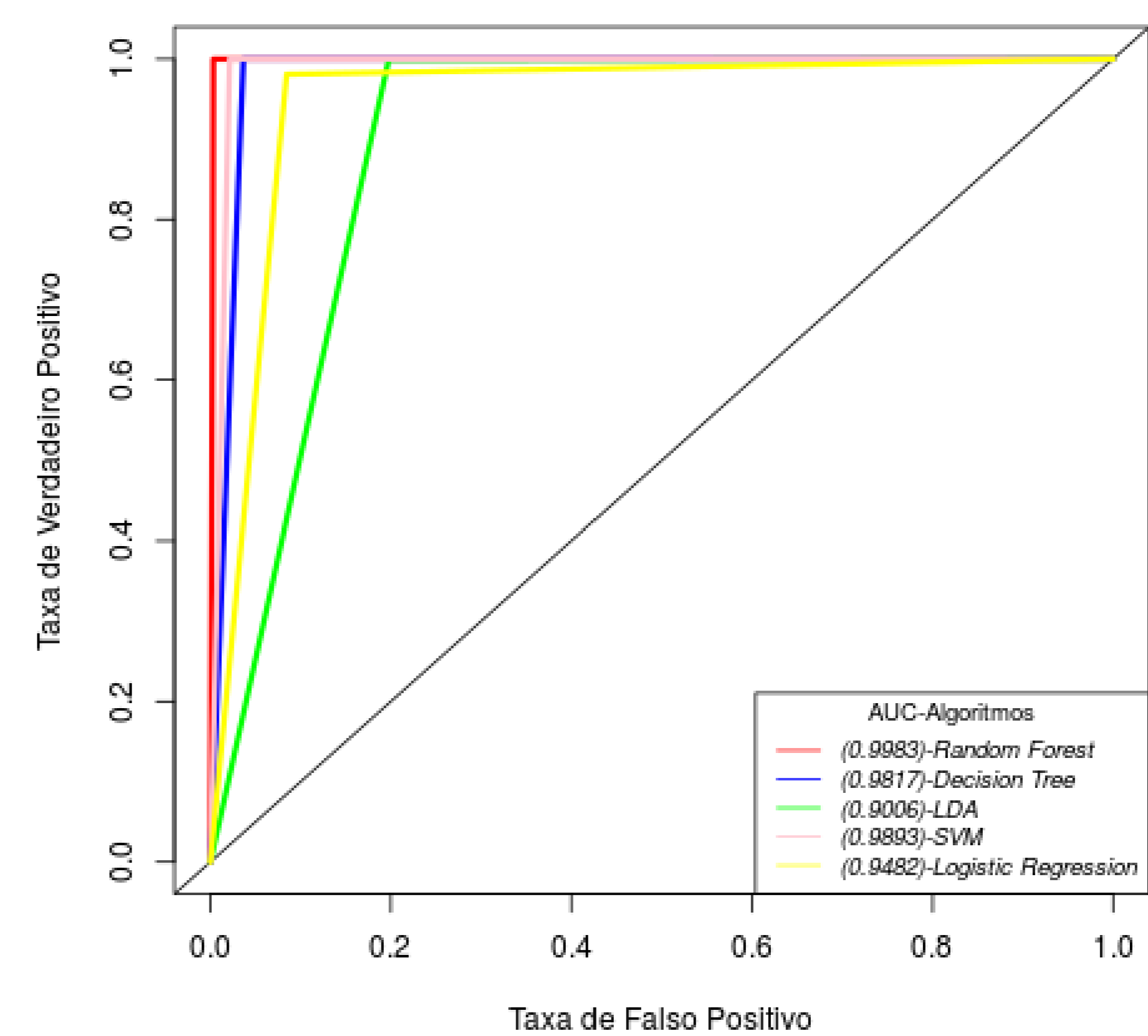


Figure 2: AUC-ROC

- O gráfico da curva *ROC*, Figura 2, apresenta um comportamento similar ao dos resultados obtidos pela medida de acurácia. Apresentando um valor de *AUC* muito próximo a 1, o *Random Forest* destaca-se entre os classificadores.

4. CONCLUSÕES

- Todos os algoritmos utilizados tiveram desempenho acurado, acima de 90%, com destaque para o classificador *Random Forest* que obteve resultados de 99% de acurácia.
- Os resultados obtidos revelam que a utilização de algoritmos de classificação pode ser uma proposta interessante para a deteção de IC.
- Essa abordagem foi discutida com o objetivo de promover a compreensão na deteção de forma autônoma de IC no domínio de Casas Inteligente à luz das pesquisas atuais na área.

Referências

- Feki M., Kawsar F., Boussard M., and Trappeniers, L. (2013). *The Internet of Things: The Next Technological Revolution*. IEEE Computer, V. 46, p.24-25