



INF1771 - Inteligência Artificial - (2018.2)
Trabalho 3 – Aprendizado de Máquina
Classificando Poker Hands

Rodrigo Pumar
Bruno Pedrazza

Poker Hands

Poker Hands - Classificador	Exemplo - 10 Atributos, 5 pares de naipe e valor (ordenados)	Possible hands	# of combinations	% of hands
Royal Straight Flush (RF)	H,1,H,10,H,11,H,12,H,13	4	480	0.000154%
Straight Flush (SF)	H,2,H,3,H,4,H,5,H,6	36	4,320	0.001385%
Four of a Kind (4)	D,8,C,8,H,8,S,8,H,12	624	74,880	0.024010%
Full House (FH)	S,2,H,2,D,7,H,7,S,7	3,744	449,280	0.144058%
Flush (F)	C,2,C,3,C,4,C,5,C,12	5,108	612,960	0.196540%
Straight (S)	C,9,D,10,H,11,D,12,S,13	10,200	1,224,000	0.392465%
Three of a Kind (3)	S,7,C,11,D,11,S,11,D,13	54,912	6,589,440	2.112845%
Two Pairs (2)	D,1,C,1,H,4,C,4,S,13	123,552	14,826,240	4.753902%
One Pair (1)	H,1,S,1,S,3,H,5,D,9	1,098,240	131,788,800	42.256903%
Only Singles(0)	S,1,S,4,S,6,C,9,C,13	1,302,540	156,304,800	50.117739%
10 Classificadores	Total	2,598,960	311,875,200	100%
		(2.5 milhões)	311 milhões	

Nossos Dados

Contagem Classificadores	25010 (25K)		1Milhão (1M)	
Classificador	Soma	Porcentagem	Soma	Porcentagem
0	12493	49.952%	501209	50.1209%
1	10599	42.379%	422498	42.2498%
2	1206	4.822%	47622	4.7622%
3	513	2.051%	21121	2.1121%
S	93	0.372%	3885	0.3885%
F	54	0.216%	1996	0.1996%
FH	36	0.144%	1424	0.1424%
4	6	0.024%	230	0.0230%
SF	5	0.020%	12	0.0012%
RF	5	0.020%	3	0.0003%
Total	25010	100,00%	1000000	100.00%

Implementação

Foi utilizado os algoritmos já implementados no WEKA, utilizando 2 tipos de algoritmos de aprendizado diferentes:

Naive Bayes

Arvore de decisão C4.5 (J48)

Foi utilizado duas tecnicas de Pre-Processamento do dados, feitas em script C#.

Foi utilizado 5-Fold Validation e dados com 25mil e 1milhão de instancias (poker hands) comparadas.



Naives Bayes

Assumimos que os atributos são independentes quando usamos Naive Bayes, oque é relativamente correto, visto que separamos cada carta em dois atributos, sendo naipe e valor, que podem ser considerados independentes.

Foi comparado o uso de estimador de densidade Kernel com distribuição normal, onde o estimador de densidade Kernel e mostrou melhor. Achamos que é porque os atributos são igualmente distribuídos, fazendo com que a distribuição normal fosse pior.

Ávores de Decisão C4.5 (J48 no WEKA)

A árvore de decisão necessita que os atributos sejam valores categóricos, como o naipe pode ser 4 (H,S,D,C) e o valor 13 (inteiros 1 até 13), esse método funciona.

Foi melhorado as configurações padrões do WEKA mudando parâmetros:

- Confiança de 0.25 para 0.5 para diminuir erro por podas, visto que a para classificar as cartas quase todos os atributos são necessários, podar a árvore sem cuidado aumentaria o erro.
- Número mínimo de instancias para cada folha de 1 (ao invés do padrão 2), visto que algumas poker hands são muito raras, poucas instancias poderiam acontecer para certas mãos.

Pré-processamento

- **Ordenamento das cartas em Valor**
- A ordem das cartas não importa no Poker. Porém, os algoritmos estavam encarando cartas iguais embaralhadas em ordem diferentes como sendo poker hands diferentes, o que não corresponde com a realidade. Como solução, decidimos ordenar as cartas pelos valores num script em C#.
- Para árvore de decisão foi duplamente útil, pois garantia que os primeiros atributos eram mais previsíveis (menores do caso de valor), sendo assim facilitava na geração de uma boa árvore de decisão.

Método	% Acertos	Tempo de build 5-fold
Naive Bayes 25K padrão	49.8481%	0.02 segundos
Naive Bayes 25K ordenado padrão	58.9924%	0.06 segundos
Naive Bayes 25K ordenado kernel	62.3790%	0.02 segundos
Árvore de decisão J48 25K padrão	53.5746%	2.9 segundos
Árvore de decisão J48 25K ordenado padrão	95.6663%	0.93 segundos
Árvore de decisão J48 25K ordenado otimizado	95.9536%	0.94 segundos

Pré-processamento

- Criação de um novo atributo se todos os Naipes eram o mesmo
- A análise dos resultados mostrou que todos os algoritmos tinham dificuldade de reconhecer que todos os naipes iguais eram uma variável importante, principalmente devido a raridade dessa poker hands e eles serem nominais.
- Foi então adicionado no pré-processamento uma nova variável, que era TRUE caso todos os atributos de naipe fossem o mesmo, e FALSE caso contrário.

Método	% Acertos	Tempo de build 5-fold	
Naive Bayes 25K ordenado kernel	62.3790%	0.02	segundos
Naive Bayes 25K ordenado kernel novo atributo	62.5990%	0.02	segundos
Árvore de decisão J48 25K ordenado otimizado	95.9536%	0.94	segundos
Árvore de decisão J48 25K ordenado otimizado novo atributo	96.4414%	0.98	segundos

Resultados Finais

Método	% Acertos	Tempo de build 5-fold (segundos)	Falso Negativo		Falso Positivo	
Naive Bayes 25K padrão	49.8481%	0.02	68	0.2719%	12475	49.8800%
Naive Bayes 25K ordenado padrão	58.9924%	0.06	2820	11.2755%	7436	29.7321%
Naive Bayes 25K ordenado kernel	62.3790%	0.02	1889	7.5530%	7520	30.0680%
Naive Bayes 25K ordenado kernel novo atributo	62.5990%	0.02	1888	7.5490%	7466	29.8521%
Naive Bayes 1M ordenado kernel novo atributo	62.3867%	1.19	71511	7.1511%	304622	30.4622%
Árvore de decisão J48 25K padrão	53.5746%	2.9	4330	17.3131%	7281	29.1124%
Árvore de decisão J48 25K ordenado padrão	95.6663%	0.93	191	0.7637%	892	3.5666%
Árvore de decisão J48 25K ordenado otimizado	95.9536%	0.94	262	1.0476%	750	2.9988%
Árvore de decisão J48 25K ordenado otimizado novo atributo	96.4414%	0.98	238	0.9516%	652	2.6070%
Árvore de decisão J48 1M ordenado otimizado novo atributo	99.9952%	157.8	2	0.0002%	46	0.0046%

Conclusão

- Aprendemos que o pré-processamento é importante para o tratamento de dados. Tivemos os exemplos prático de muito proveito com simples ordenação e criação de um atributo agregador de informação. Para problemas reais isso é útil ao invés de simplesmente tentar ajustar o algoritmo.
- No caso de árvore de decisão, a análise de raridade de algumas instancias e o modo de poda da árvore foi essencial para conseguir bons resultados.
- O quão bem um algoritmo escala tanto com pré-processamento e com número de instancias também varia, assim como seu tempo para criação de modelo. Portanto, para problemas reais onde o numero de instancias pode ser enorme e o tempo para criação do modelo crítico a análise de tempo de criação do modelo faz se essencial.