

BÁO CÁO THỰC HÀNH

Môn học: Hệ thống tìm kiếm, phát hiện và ngăn ngừa xâm nhập

Tên chủ đề: Học máy trong IDS

GVHD: Đỗ Hoàng Hiến

Nhóm: 8

1. THÔNG TIN CHUNG:

(Liệt kê tất cả các thành viên trong nhóm)

Lớp: NT204.O21.ANTT.1

STT	Họ và tên	MSSV	Email
1	Nguyễn Lê Thảo Ngọc	21521191	21521191@gm.uit.edu.vn
2	Trần Lê Minh Ngọc	21521195	21521195@gm.uit.edu.vn
3	Trần Văn Thái	21522583	21522583@gm.uit.edu.vn

2. NỘI DUNG THỰC HIỆN:¹

STT	Nội dung	Tình trạng	Trang
1	<u>Yêu cầu 1</u>	100%	2
2	<u>Yêu cầu 2.1</u>	100%	3 - 8
3	<u>Yêu cầu 2.2</u>	100%	8 - 12
4	<u>Yêu cầu 3</u>	100%	13 - 16
Điểm tự đánh giá			10/10

Phần bên dưới của báo cáo này là tài liệu báo cáo chi tiết của nhóm thực hiện.

¹ Ghi nội dung công việc, các kịch bản trong bài Thực hành

BÁO CÁO CHI TIẾT

Yêu cầu 1.1 Sinh viên tìm hiểu về tập dữ liệu KDD Cup 1999 và điền các kết quả tìm hiểu được vào form bên dưới.

TÌM HIỂU VỀ TẬP DỮ LIỆU KDD CUP 1999

Dữ liệu trong bộ dữ liệu KDD Cup 1999 là lưu lượng mạng đã được thu thập, phân tích, xử lý để lấy các thuộc tính và từ đó gán nhãn tương ứng với loại tấn công hoặc dữ liệu bình thường. Sinh viên tìm hiểu các phần sau:

1. Số nhóm tấn công: có 4 nhóm tấn công.

Kê tên các nhóm tấn công: Dos (Denial of Service), Probe, R2L (Remote to Local), U2R (User to Root)

2. Số kiểu tấn công: có 23 kiểu tấn công

Kê tên các kiểu tấn công được gán nhãn: atlantis, back, land, neptune, pod, smurf, teardrop, ipsweep, nmap, portsweep, satan, ftp_write, guess_passwd, imap, multihop, phf, spy, warezclient, warezmaster, buffer_overflow, loadmodule, perl, rootkit .

3. Mỗi instance trong tập dữ liệu KDD Cup 1999 bao gồm 41 thuộc tính, cụ thể gồm các thuộc tính:

duration, protocol_type, service, flag, src_bytes, dst_bytes, land, wrong_fragment, urgent, hot, num_failed_logins, logged_in, num_compromised, root_shell, su_attempted, num_root, num_file_creations, num_shells, num_access_files, num_outbound_cmds, is_host_login, is_guest_login, count, srv_count, error_rate, srv_error_rate, error_rate, srv_error_rate, same_srv_rate, diff_srv_rate, srv_diff_host_rate, dst_host_count, dst_host_srv_count, dst_host_same_srv_rate, dst_host_diff_srv_rate, dst_host_same_src_port_rate, dst_host_srv_diff_host_rate, dst_host_error_rate, dst_host_srv_error_rate, dst_host_error_rate, dst_host_srv_error_rate.

Yêu cầu 2.1 Sinh viên cài đặt WEKA, tìm hiểu và load một tập dữ liệu có định dạng **.arff** đơn giản có sẵn của WEKA.

Bước 1: Cài đặt WEKA

Sinh viên tham khảo tải công cụ WEKA tại đường dẫn:

<https://www.cs.waikato.ac.nz/ml/weka/downloading.html> và tham khảo các hướng dẫn cài đặt trên các hệ điều hành cụ thể

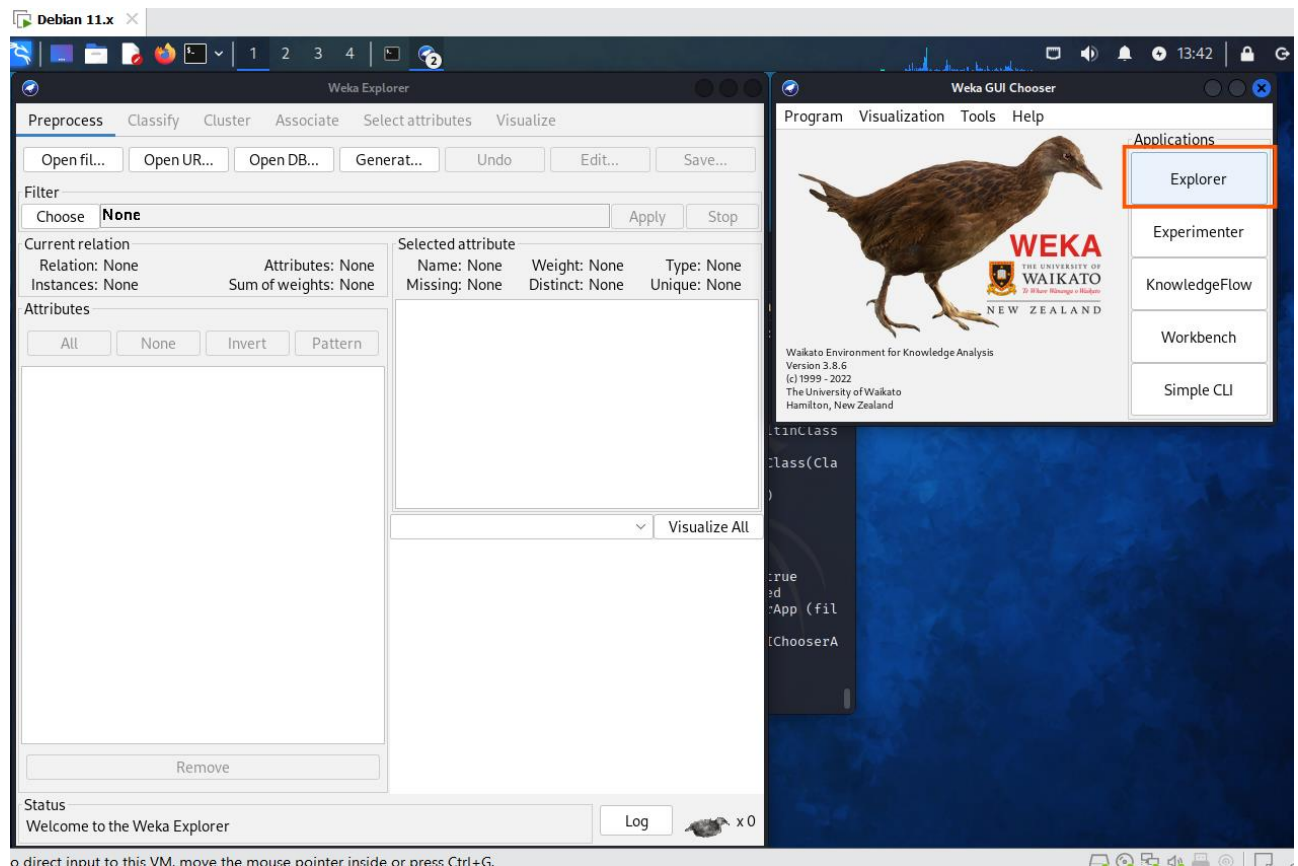
Bước 2: Chạy WEKA

Sinh viên có thể lựa chọn 2 cách chạy bằng command line hoặc sử dụng giao diện. Với lựa chọn command line, chạy lệnh sau:

```
(bun@kali) - [~/Downloads/weka-3-8-6]
$ ./weka.sh
```

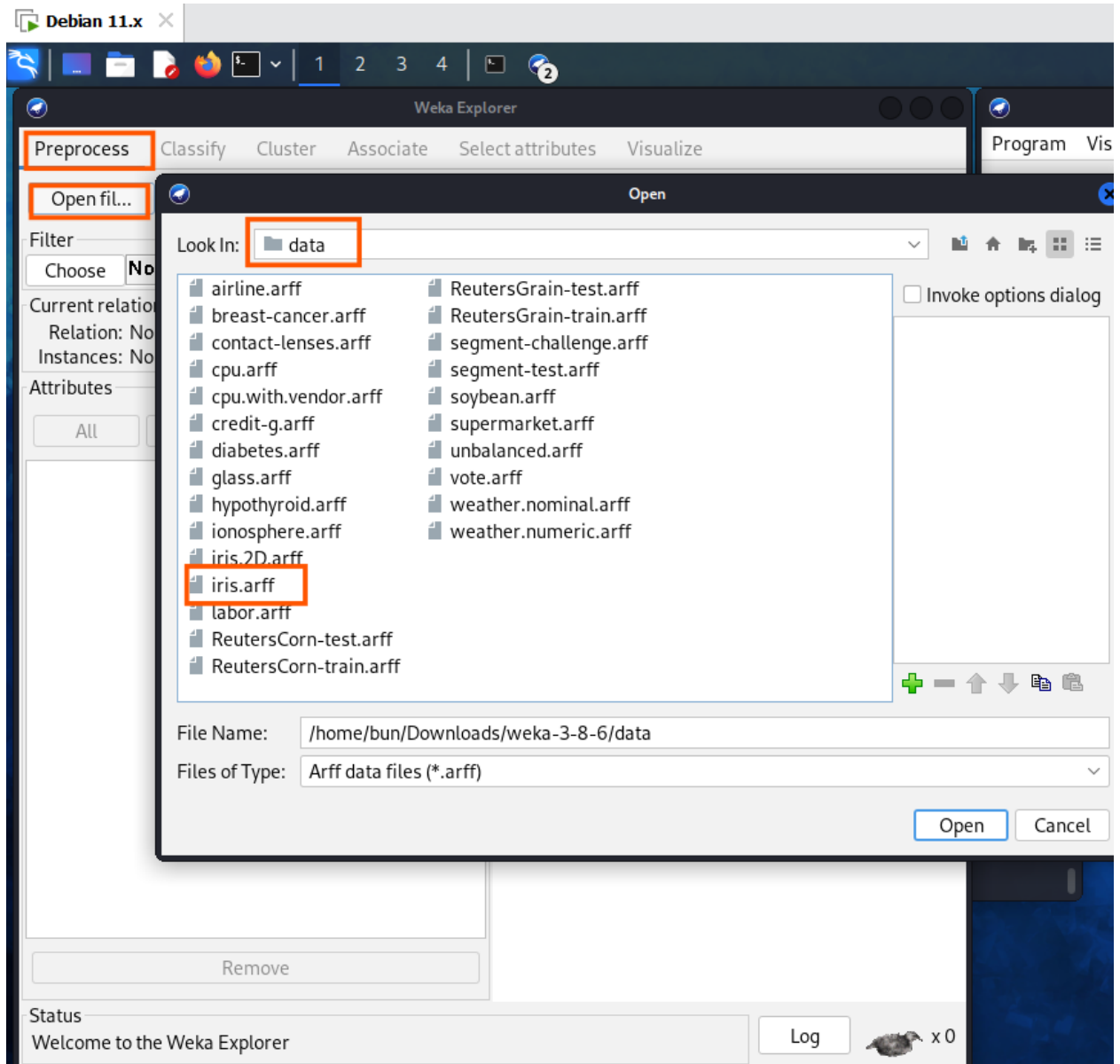
Bước 3: Trong cửa sổ Weka GUI Chooser, chọn Explorer

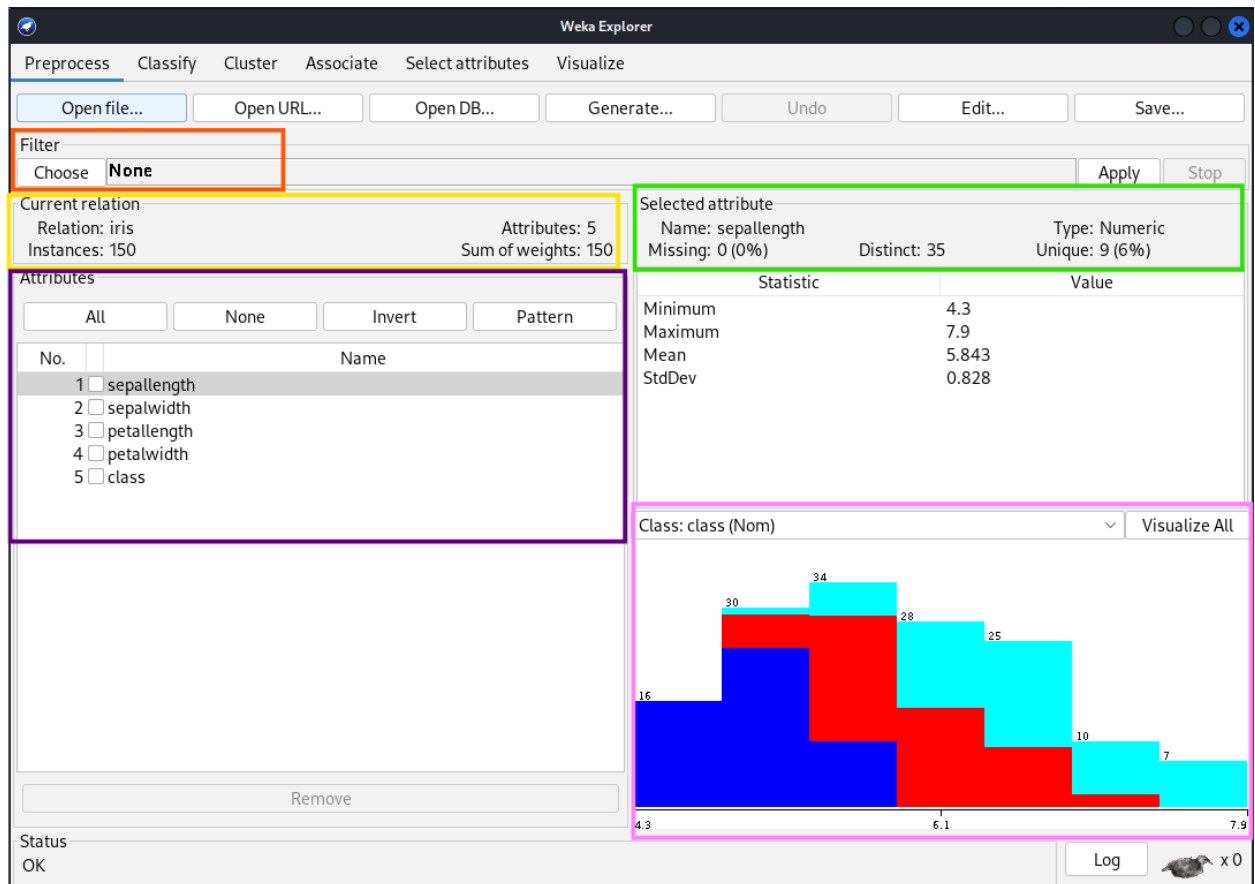
=> 1 cửa sổ mới “Weka Explorer” sẽ hiện ra



Bước 4: Load dữ liệu vào WEKA

Trong tab Preprocess, chọn Open file -> chọn folder chứa data cần mở -> chọn data muốn phân tích



Bước 5: Sinh viên quan sát và giải thích các giá trị trong tab Preprocess

Trong tab Preprocess, chúng ta thấy có 5 phần như ảnh: Filter, Attributes, Selected attribute, Class, Current relation

1. Current relation:

Phần này cho biết các thông tin tổng quát về dataset đã upload lên thông qua 4 mục:

- +Relation: tên mối quan hệ giữa các dữ liệu trong dataset
- +Instances: số lượng datapoint/record trong dataset
- +Attributes: số lượng thuộc tính (đặc trưng) trong dataset
- +Sum of weights: : tổng của các trọng số được gán cho từng mẫu dữ liệu trong tập dữ liệu.

2. Attributes

Phần này gồm 4 lựa chọn (4 tick box) và danh sách các thuộc tính trong dataset.

+4 lựa chọn gồm All, None, Invert, Pattern dùng để giúp chọn nhanh các thuộc tính cần phân tích

Nếu chọn All thì tất cả các tick box sẽ được chọn.

Nếu chọn None thì sẽ loại bỏ tất cả các chọn lựa đã thực hiện

Nếu chọn Invert thì các lựa chọn trước đó sẽ bị thay đổi ngược lại (thuộc tính chưa chọn -> được chọn; thuộc tính đã được chọn -> không chọn)

Nếu chọn Pattern, 1 cửa sổ input sẽ hiện ra, chúng ta sẽ thực hiện chọn thuộc tính mình muốn bằng cách dùng Perl 5 Regular Expression

+Danh sách lưu trữ các thuộc tính có trong dataset gồm 2 cột: No. , Name, Selection tick boxes

No. : Số thứ tự (index) của thuộc tính trong dataset

Selection tick boxes: tick box giúp chọn lựa các thuộc tính muốn phân tích.

Name: tên của thuộc tính (tên giống như trong dataset)

Mỗi khi chúng ta chọn vào 1 tên thuộc tính trong cột Name, các trường trong phần Selected attribute sẽ thay đổi nội dung ứng với thuộc tính đã chọn.

3. *Selected attributes*

Phần này sẽ hiển thị thống kê và các thông tin (type, value,...) về thuộc tính đã được chọn. Nó sẽ trích xuất đặc 1 vài đặc trưng tiêu biểu như type, missing, distinct, unique rồi hiển thị cho chúng ta thấy trong danh sách gồm:

+Name: tên của thuộc tính

+Type: dạng dữ liệu của thuộc tính (ví dụ như: numeric, categorical)

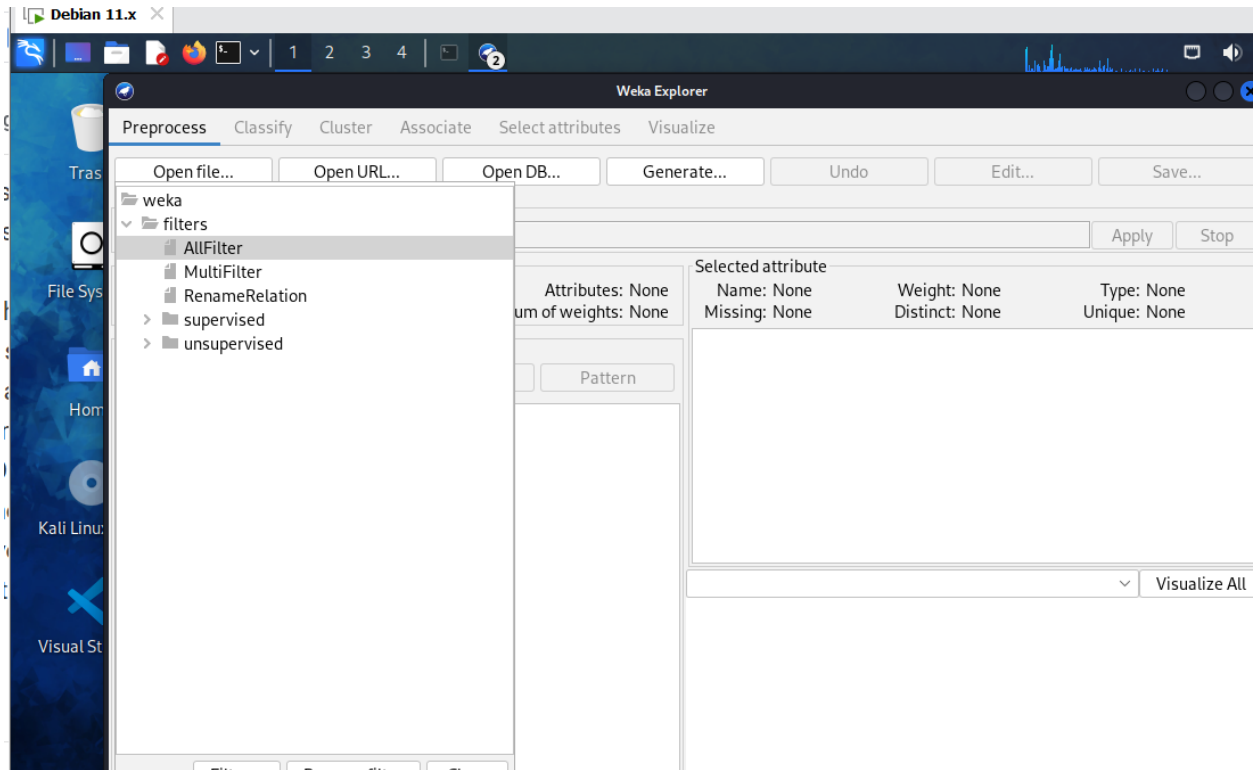
+Missing: số lượng và tỉ lệ dữ liệu bị khuyết của thuộc tính

+Distinct: Số lượng giá trị khác nhau trong thuộc tính này

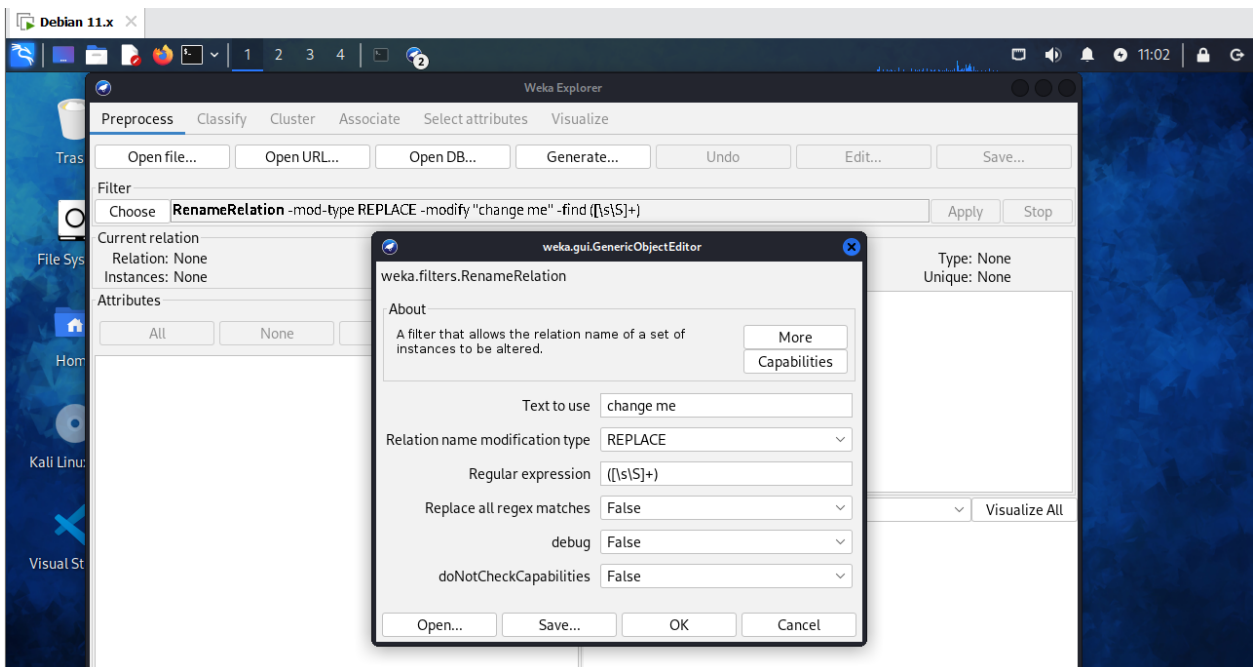
+Unique: Số lượng hoặc tỉ lệ giá trị duy nhất mà thuộc tính này có trong dataset

4. *Filter*

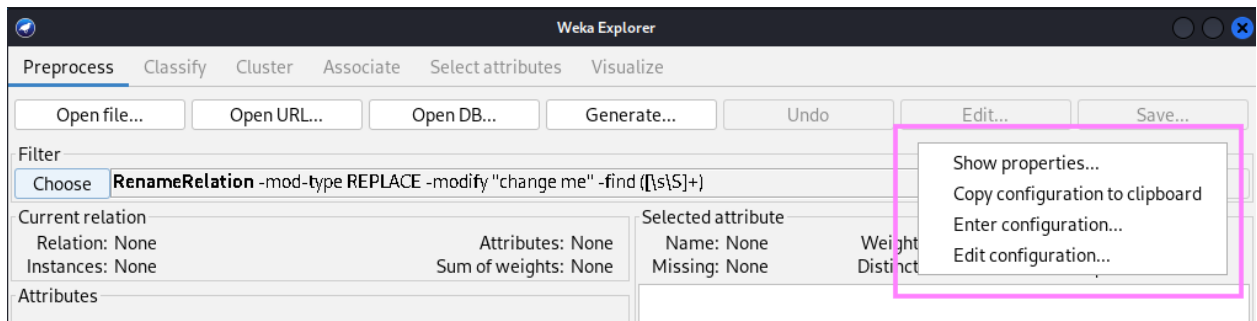
Phần này dùng để thiết lập các bộ lọc theo yêu cầu người dùng. Khi nhấn vào nút Choose, một danh sách các filter sẽ hiển ra để người dùng lựa chọn. Sau khi đã chọn 1 loại filter, tên và các tùy chọn ứng với filter đó sẽ được hiển thị trong trường kế bên nút Choose.



Nếu nhấn trái chuột vào trường này thì 1 hộp thoại GenericObjectEditor (dùng để cấu hình filter đã chọn) sẽ hiện ra.



Nếu nhấn phải chuột vào trường này thì 1 menu xuất hiện. Menu này gồm các lựa chọn như hiển thị thông số của filter, chỉnh sửa cấu hình hay sao chép cấu hình hiện tại vào clipboard



Sau khi chọn và hoàn tất cấu hình cho filter đã chọn, chúng ta nhấn nút Apply để áp dụng vào dữ liệu.

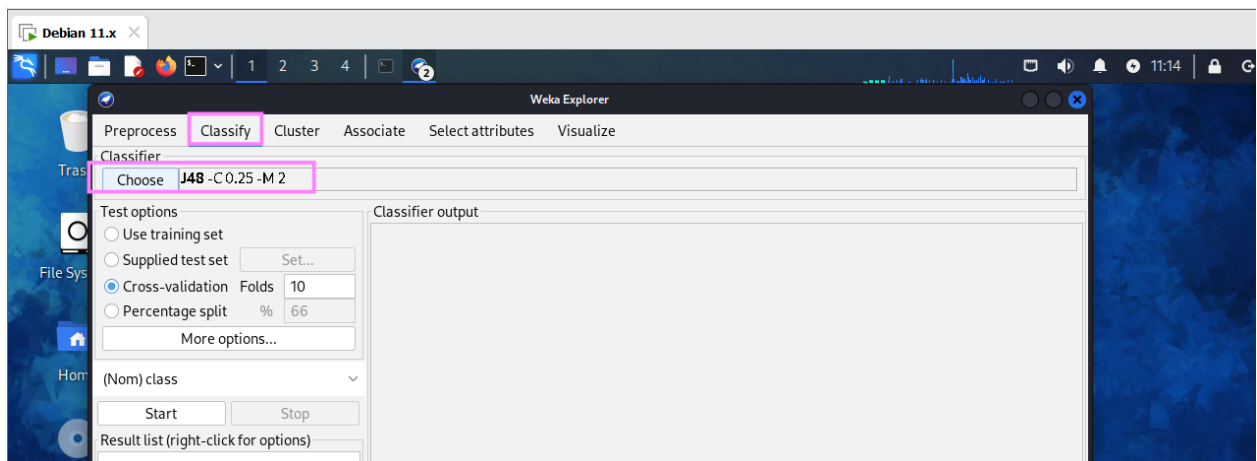
Yêu cầu 2.2 Sinh viên lựa chọn 01 bộ phân lớp (classifier) bất kỳ và thực hiện khai thác trên tập dữ liệu đã chọn ở trên. Trình bày và giải thích kết quả.

Yêu cầu: Giải thích test option lựa chọn và ý nghĩa của các thông số kết quả đầu ra sau khi chạy bộ phân lớp

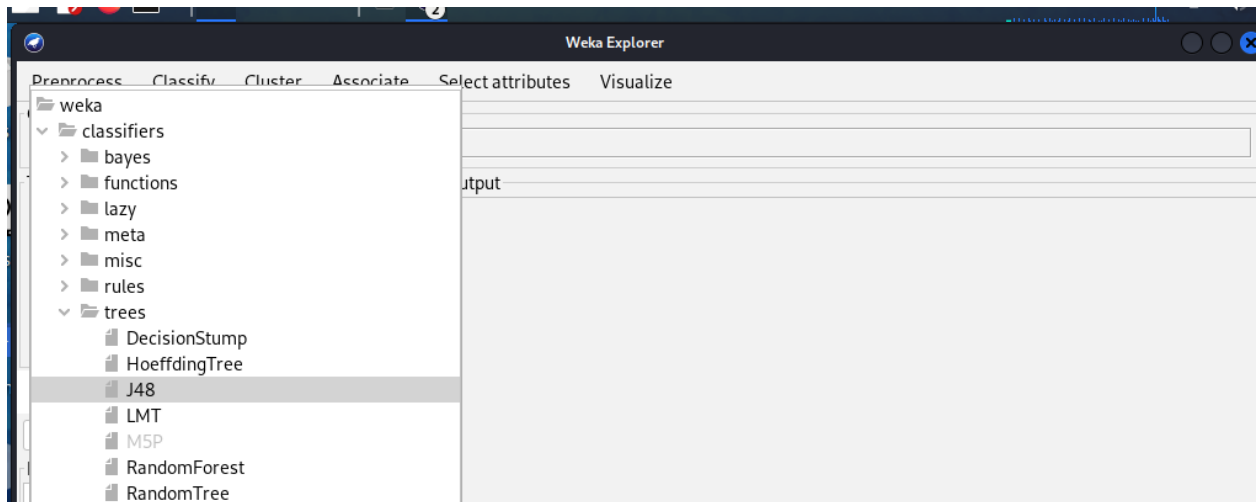
Nhóm chúng em sẽ chọn tập dữ liệu sẵn có trong WEKA là iris.arff để khai thác. Sau đó tiến hành chọn bộ phân lớp.

Bước 1: Lựa chọn bộ phân lớp cho tập dữ liệu

Trong cửa sổ WEKA Explorer, chọn tab **Classify**

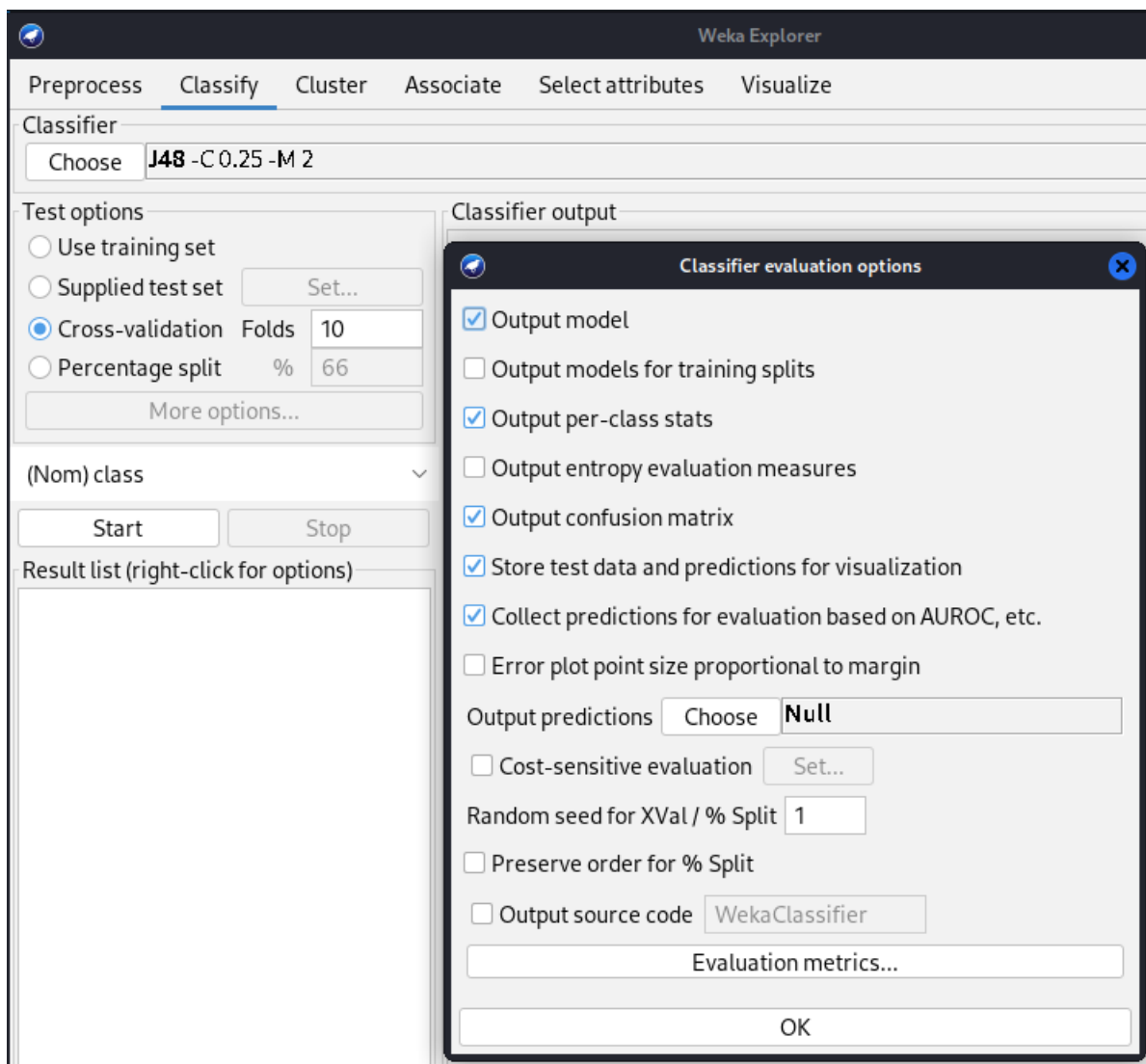


Trong khung lựa chọn **Classifier**, nhấp chọn **Choose** để hiển thị danh sách các bộ phân lớp được WEKA hỗ trợ. Tùy vào tập dữ liệu được load lên mà chỉ có những bộ phân lớp phù hợp có thể chạy trên đó mới khả dụng.



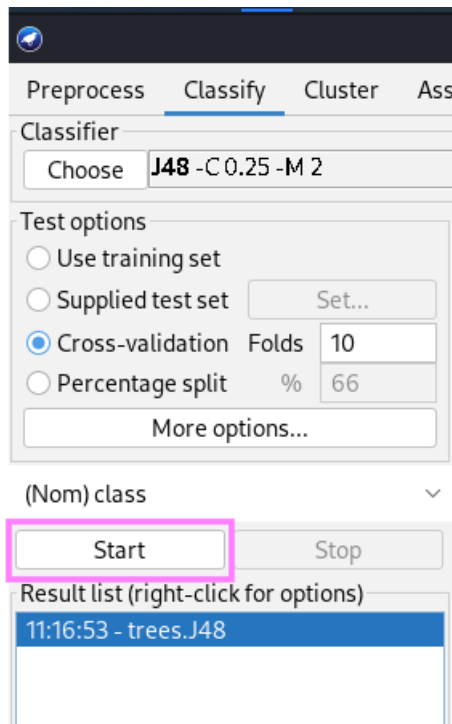
Bước 2: Lựa chọn Test options

Đây là các tùy chọn được WEKA hỗ trợ để định nghĩa 2 tập dữ liệu huấn luyện và kiểm tra.



Bước 3: Chạy bộ phân lớp và quan sát kết quả

Nhấn chọn **Start** để bắt đầu chạy bộ phân lớp.

**Bước 4: Quan sát và giải thích kết quả**

Một danh sách thống kê mức độ chính xác của bộ phân lớp hiện ra

Classifier output

Size of the tree : 9

Time taken to build model: 0.03 seconds

=== Stratified cross-validation ===

=== Summary ===

Metric	Value	Percentage
Correctly Classified Instances	144	96 %
Incorrectly Classified Instances	6	4 %
Kappa statistic	0.94	
Mean absolute error	0.035	
Root mean squared error	0.1586	
Relative absolute error	7.8705 %	
Root relative squared error	33.6353 %	
Total Number of Instances	150	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.980	0.000	1.000	0.980	0.990	0.985	0.990	0.987	Iris-setosa
	0.940	0.030	0.940	0.940	0.940	0.910	0.952	0.880	Iris-versicolor
	0.960	0.030	0.941	0.960	0.950	0.925	0.961	0.905	Iris-virginica
Weighted Avg.	0.960	0.020	0.960	0.960	0.960	0.940	0.968	0.924	

=== Confusion Matrix ===

```

a b c <-- classified as
49 1 0 | a = Iris-setosa
0 47 3 | b = Iris-versicolor
0 2 48 | c = Iris-virginica
  
```

- Size of the tree (Kích thước của cây): Đây là số lượng nút (node) trong cây quyết định sau quá trình xây dựng. Trong trường hợp này, cây có 9 nút.
 - Time taken to build model (Thời gian xây dựng mô hình): Đây là thời gian để xây dựng mô hình, được tính bằng giây. Trong trường hợp này, thời gian xây dựng mô hình là 0.03 giây.
 - Stratified cross-validation (Kiểm định chéo theo phân lớp): Đây là phương pháp kiểm định mô hình được sử dụng để đánh giá hiệu suất của mô hình trên tập dữ liệu. Kết quả của quá trình kiểm định chéo được hiển thị như sau:
 - + Correctly Classified Instances (Số lượng mẫu được phân loại chính xác): Đây là số lượng mẫu dữ liệu được phân loại đúng bởi mô hình. Trong trường hợp này, có 144 mẫu được phân loại đúng (96%).
 - + Incorrectly Classified Instances (Số lượng mẫu được phân loại sai): Đây là số lượng mẫu dữ liệu bị phân loại sai bởi mô hình. Trong trường hợp này, có 6 mẫu được phân loại sai (4%).
 - + Kappa statistic (Thống kê Kappa): Đây là một thống kê sử dụng để đo độ đồng nhất giữa các dự đoán của mô hình và nhãn thực tế. Giá trị Kappa bằng 0.94 xấp xỉ 1 cho thấy mô hình có độ chính xác cao.
 - + Mean absolute error (Sai số trung bình tuyệt đối): Đây là giá trị trung bình của sai số tuyệt đối giữa dự đoán và nhãn thực tế.
 - + Root mean squared error (Sai số bình phương trung bình căn): Đây là sai số trung bình của dự đoán so với nhãn thực tế, tính bằng cách lấy căn bậc hai của sai số bình phương trung bình.
 - + Relative absolute error (Sai số tuyệt đối tương đối): Đây là tỷ lệ giữa sai số tuyệt đối trung bình và giá trị trung bình của nhãn thực tế.
 - + Root relative squared error (Sai số bình phương tương đối căn): Đây là tỷ lệ giữa sai số bình phương trung bình căn và giá trị trung bình của nhãn thực tế.
 - + Total Number of Instances (Tổng số mẫu dữ liệu): Đây là tổng số mẫu dữ liệu trong tập dữ liệu. Trong trường hợp này, có tổng cộng 150 mẫu.
 - Detailed Accuracy By Class (Độ chính xác chi tiết theo lớp): Phân tích chi tiết hơn về độ chính xác trong dự đoán theo từng lớp riêng biệt trong bộ phân loại. Nó gồm các phương pháp đánh giá mô hình machine learning như độ nhạy (recall), độ chính xác dự đoán (precision), F-Measure, MCC, ROC Area, PRC Area.
- +TP rate: tỉ lệ true positive khi dự đoán
- True positives: Các điểm Positive thực được phân lớp đúng là Positive
- +FP rate: tỉ lệ false positive khi dự đoán

False positives: Các điểm Negative thực được phân lớp sai là Positive

+ Precision (Độ chính xác): Precision được định nghĩa là tỉ lệ số điểm true positive trong số những điểm được phân loại là *positive* (TP + FP). Thể hiện sự chuẩn xác của việc phát hiện các điểm Positive. Số này càng cao thì model nhận các điểm Positive càng chuẩn.

+ Recall (Độ phủ): được định nghĩa là tỉ lệ số điểm true positive trong số những điểm thực sự là positive (TP + FN). Thể hiện khả năng phát hiện tất cả các positive, tỷ lệ này càng cao thì cho thấy khả năng bỏ sót các điểm Positive là thấp

+ F1-score: F1-score là một chỉ số kết hợp giữa precision và recall để đánh giá hiệu suất tổng thể của mô hình.

- Confusion Matrix (Ma trận nhầm lẫn): Đây là confusion matrix, hiển thị số lượng mẫu thực tế được phân loại vào từng lớp, so sánh với lớp dự đoán của mô hình. Mỗi cột là 1 lớp dự đoán, hàng là số lượng mẫu dự đoán thuộc lớp đó.

Nó giúp đánh giá sự khác biệt giữa dự đoán và nhãn thực tế.

Yêu cầu 3.1 Sinh viên lựa chọn 01 bộ phân lớp bất kỳ và thực hiện khai thác trên tập dữ liệu KDD Cup 1999. Giải thích và đánh kết quả.

Giải thích các test options:

- Use training set: Dataset dùng để train được dùng test model
- Supplied test set: Dùng dataset độc lập với data train để test model.
- Cross-validation: Với tham số K fold, nó sẽ chia dataset thành K set con, dùng K-1 set để train và dùng set còn lại để test. Lặp lại quá trình này cho từng set còn lại cho đến khi hết K set đó.
- Percentage split: Với tham số K%, nghĩa là nó sẽ dùng K% dataset để train, số còn lại dùng để test.

Classifier được chọn: RandomForest.

Nguyên lý hoạt động là bằng cách kết hợp các dự đoán của cây quyết định nhỏ hơn để tổng hợp vào tạo ra dự đoán cuối cùng. Và test model bằng option Cross-validation với folds là 10.

Lý do chọn:

- Kết quả output có độ chính xác cao.
- Không yêu cầu quá nhiều việc tiền xử lý data.
- Xử lý được đa dạng dữ liệu như dạng số, dạng danh mục,...

Kết quả đánh giá:

- Với bộ dataset đưa vào có tất cả 494020 mẫu, sau khi train, kết quả thu được khitest model cho ra kết quả như sau:

Weka Explorer

Preprocess **Classify** Cluster Associate Select attributes Visualize

Classifier: Choose **RandomForest** -P 100 -I 100 -num-slots 1 -K 0 -M 1.0 -V 0.001 -S 1

Test options

☐ Use training set

☐ Supplied test set Set...

☒ Cross-validation Folds **10**

☐ Percentage split % **66**

More options...

(Nom) result **Start** **Stop**

Result list (right-click for options)

14:42:33 - rules.ZeroR

14:42:43 - trees.RandomForest

Classifier output

=== Summary ===

Correctly Classified Instances	493523	99.9804 %
Incorrectly Classified Instances	97	0.0196 %
Kappa statistic	0.9997	
Mean absolute error	0.0001	
Root mean squared error	0.004	
Relative absolute error	0.1056 %	
Root relative squared error	2.5217 %	
Total Number of Instances	494020	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	1.000	0.000	0.999	1.000	1.000	1.000	1.000	1.000	normal.
	0.867	0.000	0.788	0.867	0.825	0.826	1.000	0.930	buffer_overflow.
	0.222	0.000	1.000	0.222	0.364	0.471	1.000	0.691	loadmodule.
	0.667	0.000	1.000	0.667	0.800	0.816	1.000	1.000	perl.
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	neptune.
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	smurf.
	0.962	0.000	1.000	0.962	0.981	0.981	1.000	0.996	guess_passwd.
	0.996	0.000	0.996	0.996	0.996	0.996	1.000	1.000	pod.
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	teardrop.
	0.996	0.000	0.998	0.996	0.997	0.997	1.000	0.999	portswEEP.
	0.992	0.000	0.996	0.992	0.994	0.994	1.000	1.000	ipswEEP.
	0.905	0.000	0.950	0.905	0.927	0.927	1.000	0.979	land.
	0.500	0.000	1.000	0.500	0.667	0.707	1.000	0.604	ftp_write.
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	back.
	0.917	0.000	1.000	0.917	0.957	0.957	1.000	1.000	imap.
	0.991	0.000	1.000	0.991	0.995	0.995	1.000	0.999	satAn.
	0.750	0.000	1.000	0.750	0.857	0.866	1.000	1.000	phf.
	0.974	0.000	1.000	0.974	0.987	0.987	1.000	0.996	nmAP.
	0.429	0.000	0.600	0.429	0.500	0.507	1.000	0.602	multihop.
	0.800	0.000	0.889	0.800	0.842	0.843	0.975	0.899	wareZmaster.
	0.993	0.000	0.993	0.993	0.993	0.993	1.000	1.000	wareZclient.
	0.000	0.000	?	0.000	?	?	1.000	0.450	spy.
	0.100	0.000	0.500	0.100	0.167	0.224	1.000	0.222	rootkit.

Status
OK

	0.750	0.000	1.000	0.750	0.857	0.866	1.000	1.000	phf.
	0.974	0.000	1.000	0.974	0.987	0.987	1.000	0.996	nmap.
	0.429	0.000	0.600	0.429	0.500	0.507	1.000	0.602	multihop.
	0.800	0.000	0.889	0.800	0.842	0.843	0.975	0.899	warezmaster.
	0.993	0.000	0.993	0.993	0.993	0.993	1.000	1.000	warezclient.
	0.000	0.000	?	0.000	?	?	1.000	0.450	spy.
	0.100	0.000	0.500	0.100	0.167	0.224	1.000	0.222	rootkit.
Weighted Avg.	1.000	0.000	?	1.000	?	?	1.000	1.000	

=== Confusion Matrix ===

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	<-- classifi
97266	2	0	0	0	1	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	5	0	1	a = nor
4	26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	b = buf
4	3	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	c = loa
1	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	d = per
1	0	0	0	107199	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	e = nep
0	0	0	0	0	280790	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	f = smu
2	0	0	0	0	0	51	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	g = que
1	0	0	0	0	0	0	263	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	h = pod
0	0	0	0	0	0	0	0	979	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	i = tea
2	0	0	0	0	1	0	0	0	0	1036	1	0	0	0	0	0	0	0	0	0	0	0	0	j = por
9	0	0	0	0	0	0	0	0	0	1	1237	0	0	0	0	0	0	0	0	0	0	0	0	k = ips
0	0	0	0	2	0	0	0	0	0	0	0	19	0	0	0	0	0	0	0	0	0	0	0	l = lan
2	1	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	1	0	0	m = ftp
0	0	0	0	0	0	0	0	0	0	0	0	0	2203	0	0	0	0	0	0	0	0	0	0	n = bac
1	0	0	0	0	0	0	0	0	0	0	0	0	0	11	0	0	0	0	0	0	0	0	0	o = ima
13	0	0	0	2	0	0	0	0	0	0	0	0	0	0	1574	0	0	0	0	0	0	0	0	p = eat
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	q = phf
2	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	225	0	0	0	0	0	0	r = nma
0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	2	1	0	0	0	s = mul
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	16	0	0	0	0	t = war
7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1013	0	0	0	u = war
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	v = spy
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	w = roo

Hình cho thấy model có mức độ phát hiện chính xác rất cao với khoảng ~99%, với các chỉ số khá tốt

- Kappa statistic: cho biết mức độ chính xác của lớp phân loại, từ -1 đến 1, càng gần 1 thì càng tốt
- Mean Absolute Error: độ lỗi trung bình giữa các dự đoán và giá trị thực tế.
- Root Mean Absolute Error: độ lỗi trung bình giữa các dự đoán và giá trị thực tế.
- Relative absolute error: độ lỗi tương đối giữa dự đoán và giá trị thực tế

Weka Explorer

Preprocess **Classify** Cluster Associate Select attributes Visualize

Classifier: Choose **RandomTree -K 0-M 1.0-V 0.001-S 1**

Test options

- ☐ Use training set
- ☐ Supplied test set
- ☒ Cross-validation Folds: **10**
- ☐ Percentage split %: **66**

More options...

(Nom) class

Start Stop

Result list (right-click for options)

- 14:42:43 - trees.RandomForest
- 15:12:51 - rules.ZeroR**
- 15:33:43 - trees.RandomTree

Classifier output

```
=====  
class  
Test mode: 10-fold cross-validation  
==== Classifier model (full training set) ====  
ZeroR predicts class value: Iris-setosa  
Time taken to build model: 0 seconds  
==== Stratified cross-validation ====  
==== Summary ====  
Correctly Classified Instances      50           33.3333 %  
Incorrectly Classified Instances    100           66.6667 %  
Kappa statistic                     0  
Mean absolute error                 0.4444  
Root mean squared error             0.4714  
Relative absolute error             100 %  
Root relative squared error         100 %  
Total Number of Instances          150  
  
==== Detailed Accuracy By Class ====  
      TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class  
1.000  1.000  0.333  1.000  0.500  7  0.500  0.333  Iris-setosa  
0.000  0.000  ?  0.000  ?  7  0.500  0.333  Iris-versicolor  
0.000  0.000  ?  0.000  ?  7  0.500  0.333  Iris-virginica  
Weighted Avg.  0.333  0.333  ?  0.333  ?  7  0.500  0.333  
  
==== Confusion Matrix ====  
a b c <-- classified as  
50 0 0 | a = Iris-setosa  
50 0 0 | b = Iris-versicolor  
50 0 0 | c = Iris-virginica
```

Status: OK

Log x0