



SDAIA
الهيئة السعودية للبيانات
والذكاء الاصطناعي
Saudi Data & AI Authority

Predicting qualified employees for promotion

By:
Shahad Almutairi
Sahar Alharthi
Ghaida Alaqeel

Instructor:
Dr. Mejdal Alqahtani


INTRODUCTION

الرئيسية / تحقيقات وتقارير

الثلاثين، 17 جمادى الأولى 1433 هـ - 9 أبريل 2012م العدد 15995

د. الهيجان: تقسيم الموظفين إلى فئات ولكل فئة معاملة خاصة

موظفون فاتهم «قطار الترقية» وينتظرون 12 عاماً لتحقيق الحلم



د.عبدالرحمن الهيجان
www.alriyadh.com

إعداد: عيبر البراهيم

INTRODUCTION

التعيين حتى نهاية الخدمة

وهنا أهم الخدمات

- سجل الموظف**
يمكن الموظف من الاطلاع على بياناته، ومؤهلاته، وخبراته، وناسخه الوظيفي
- الترقية**
تتيح للموظف طلب الترقية، وتحديد رغبته، والدخول في المفاضلة إلكترونياً، ومتابعتها
- الشواغر**
تقدم الإعلان عن الشواغر الوظيفية، تمكن الموظف من الاطلاع عليها، والتقدم لها إذا ناسبته
- التظلمات**
تتيح للموظف إن وقعت عليه مظلمة، أن يرفع طلب التظلم للوزارة إلكترونياً، ويتابعه
- مطابقة البيانات**
تمكن الجهات من مطابقة البيانات مع وزارة الخدمة المدنية، لعرضها وتصحيحها ومعالجتها خلال فترة وجيزة
- التعاقد**
تسهل عمليات التعاقد مع الموظفين وتسريعها من خلال التحقق من المرشحين لرؤية إمكانية التعاقد معهم
- التوطين**

فكرة
fekora.com



The goal:

Predicting employees who are qualified for promotion.

DATA STRUCTURE

Columns: 13

Rows: 54808

	employee_id	department	region	education	gender	recruitment_channel	no_of_trainings	age	previous_year_rating	length_of_service	awards_won?	avg_training_score	is_promoted
0	65438	Sales & Marketing	region_7	Master's & above	f	sourcing	1	35	5.0	8	0	49	0
1	65141	Operations	region_22	Bachelor's	m	other	1	30	5.0	4	0	60	0
2	7513	Sales & Marketing	region_19	Bachelor's	m	sourcing	1	34	3.0	7	0	50	0
3	2542	Sales & Marketing	region_23	Bachelor's	m	other	2	39	1.0	10	0	50	0
4	48945	Technology	region_26	Bachelor's	m	other	1	45	3.0	2	0	73	0
4													

DATA CLEANING

NULLS VALUE - DUPLICATED - MANAGING COLUMNS

After Cleaning...

FEATURES

DEPARTMENT

EDUCATION

GENDER

NO OF
TRAININGS

AGE

PREVIOUS YEAR
RATING

LENGTH OF
SERVICE

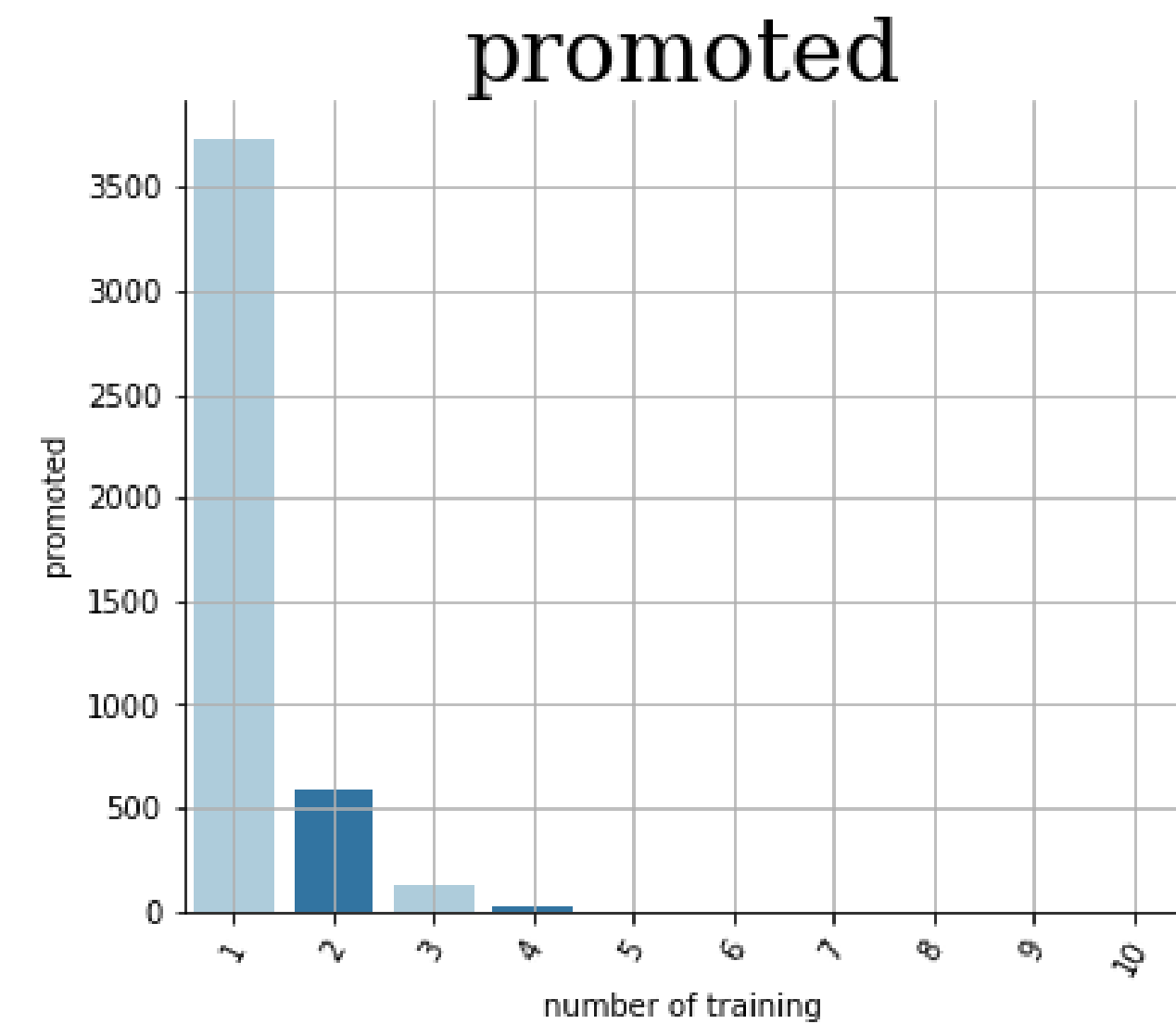
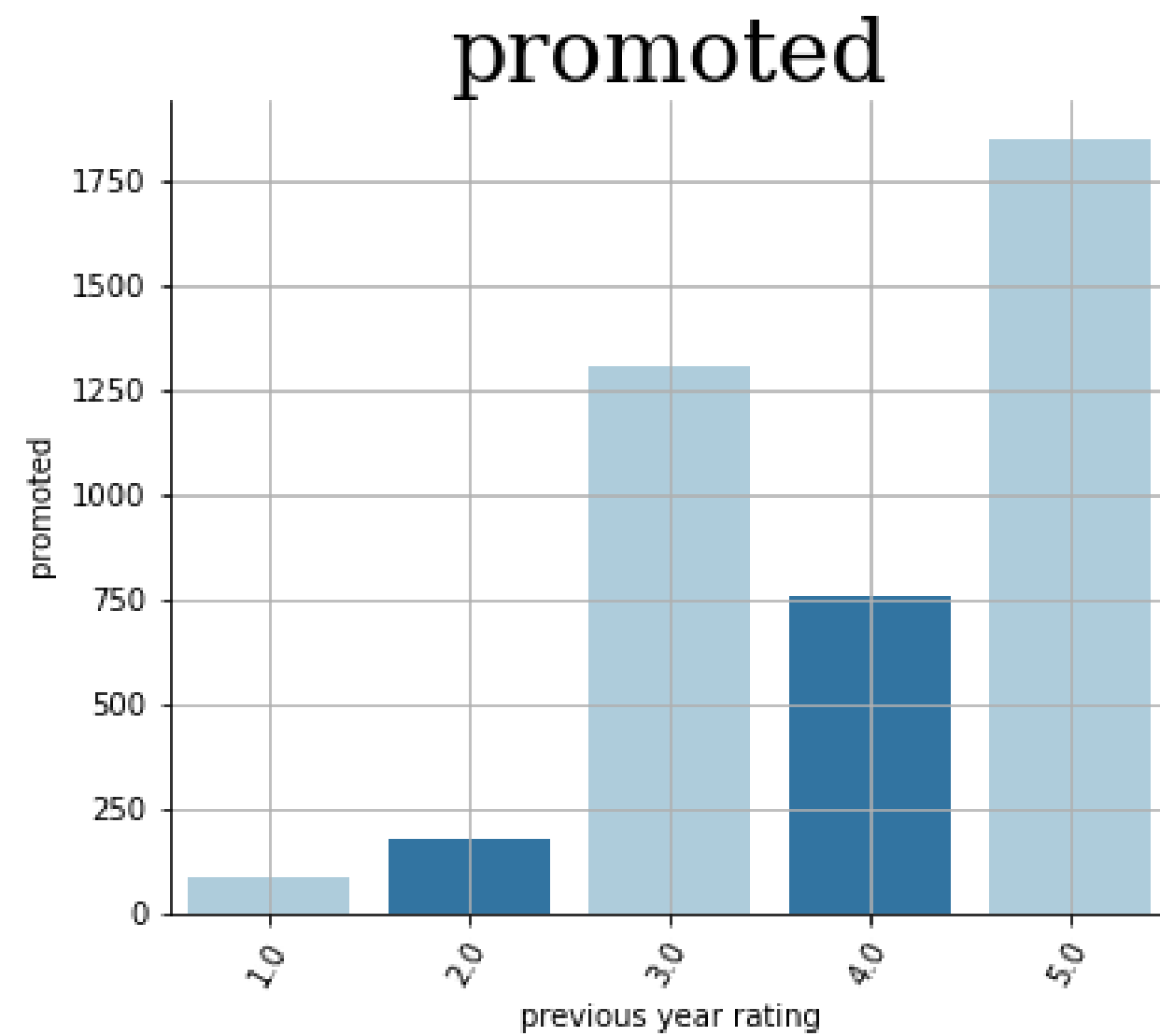
AWARDS WON

AVREGE
TRAINING
SCORE

TARGET

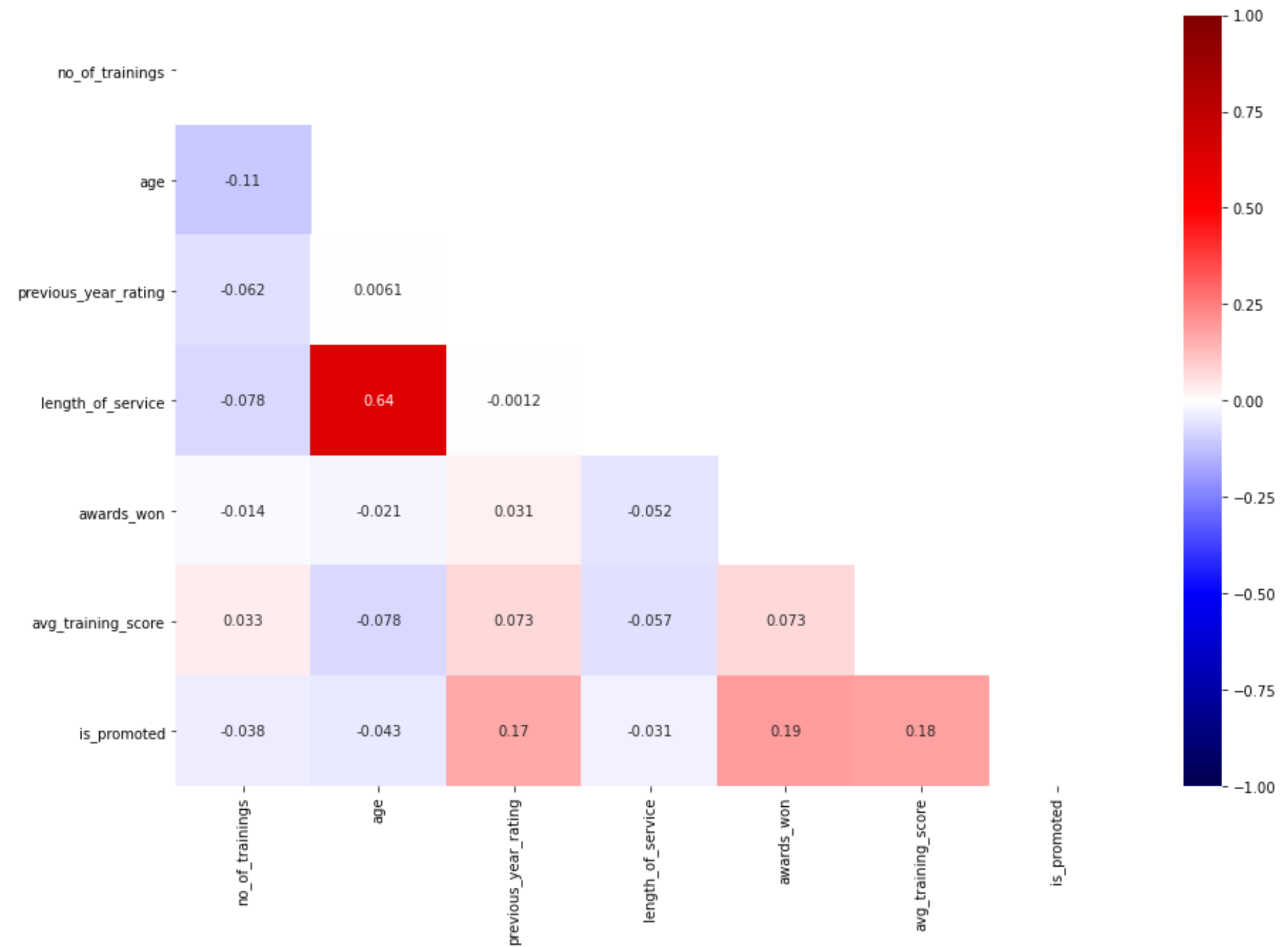
IS PROMOTED

DATA ANALYSIS



DATA ANALYSIS

Correlation Heatmap



Classification Models

**LOGISTIC
REGRESSION**

**DECISION
TREE**

**RANDOM
FOREST**

**SUPPORT
VECTOR
MACHINE**

KNN

**GRADIENT
BOOSTING**

**ADAPTTIVE
BOOSTING**

IMPALANCE

BEFOUR

AFTER

0 - 32867

1- 3618

SMOTE



0 - 32867

1- 32867

KNN

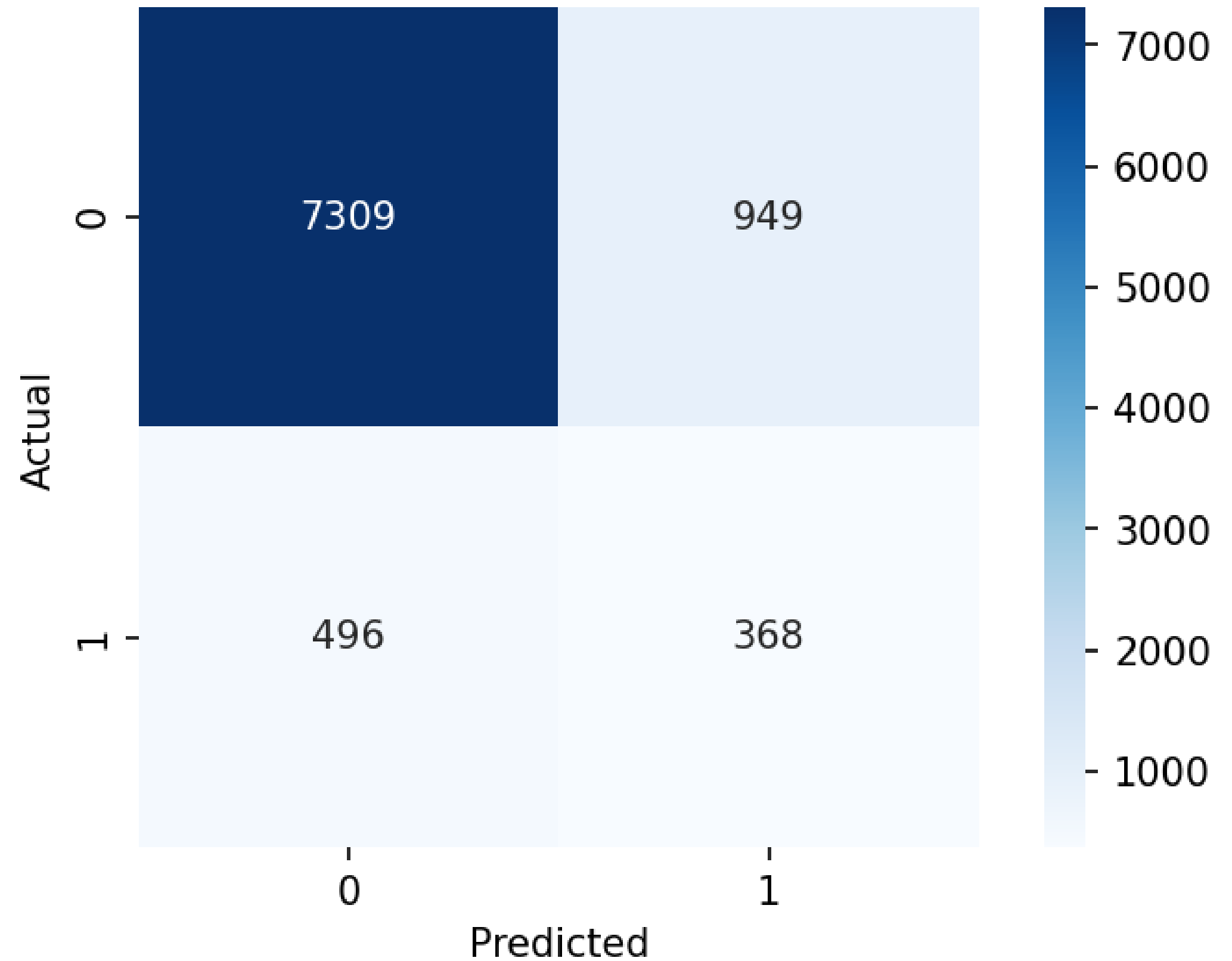
Train

90%

Test

84%

KNN regression confusion matrix



LOGISTIC REGRESSION

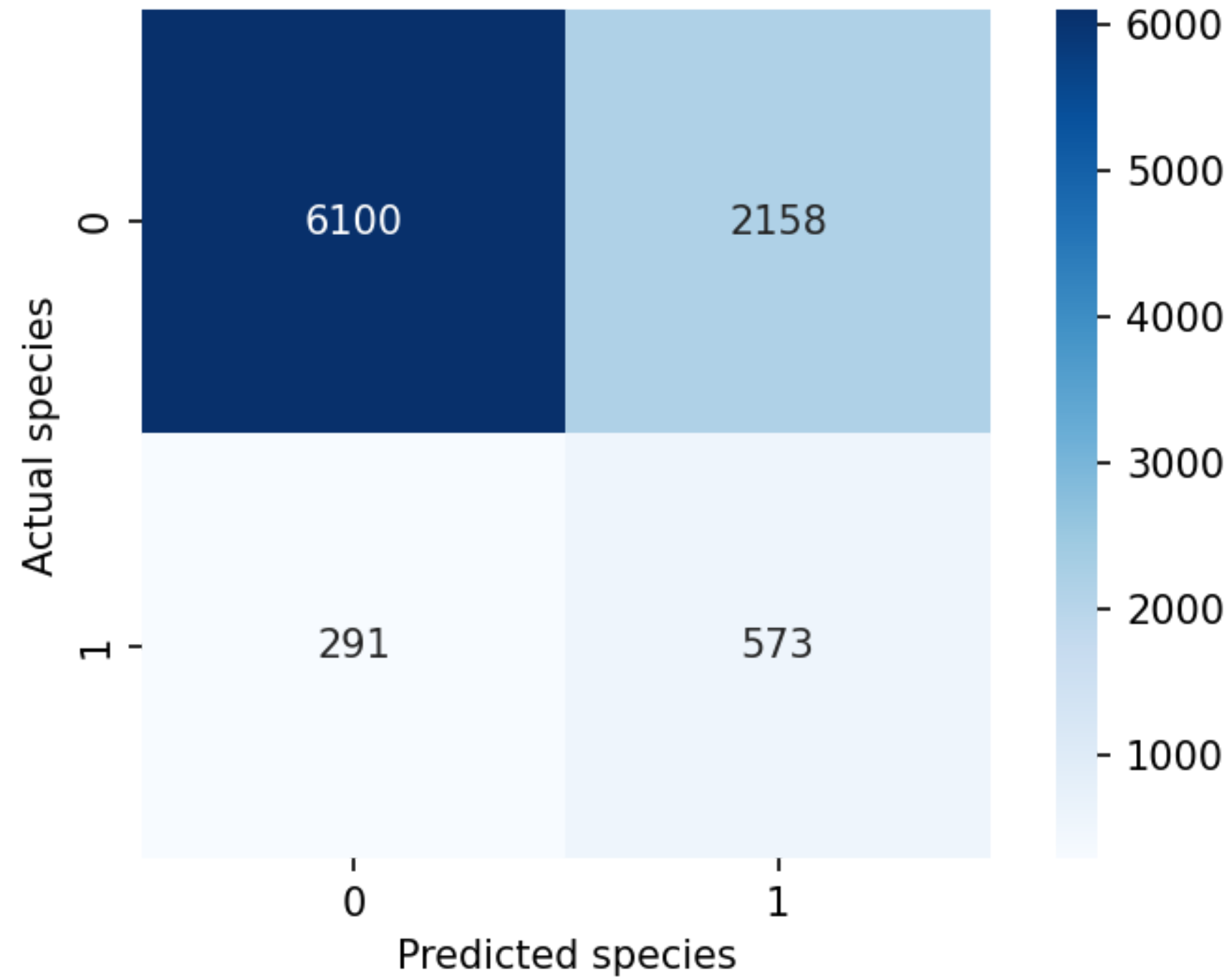
Train

73%

Test

71%

Logistic regression confusion matrix



SUPPORT VECTOR MACHINE

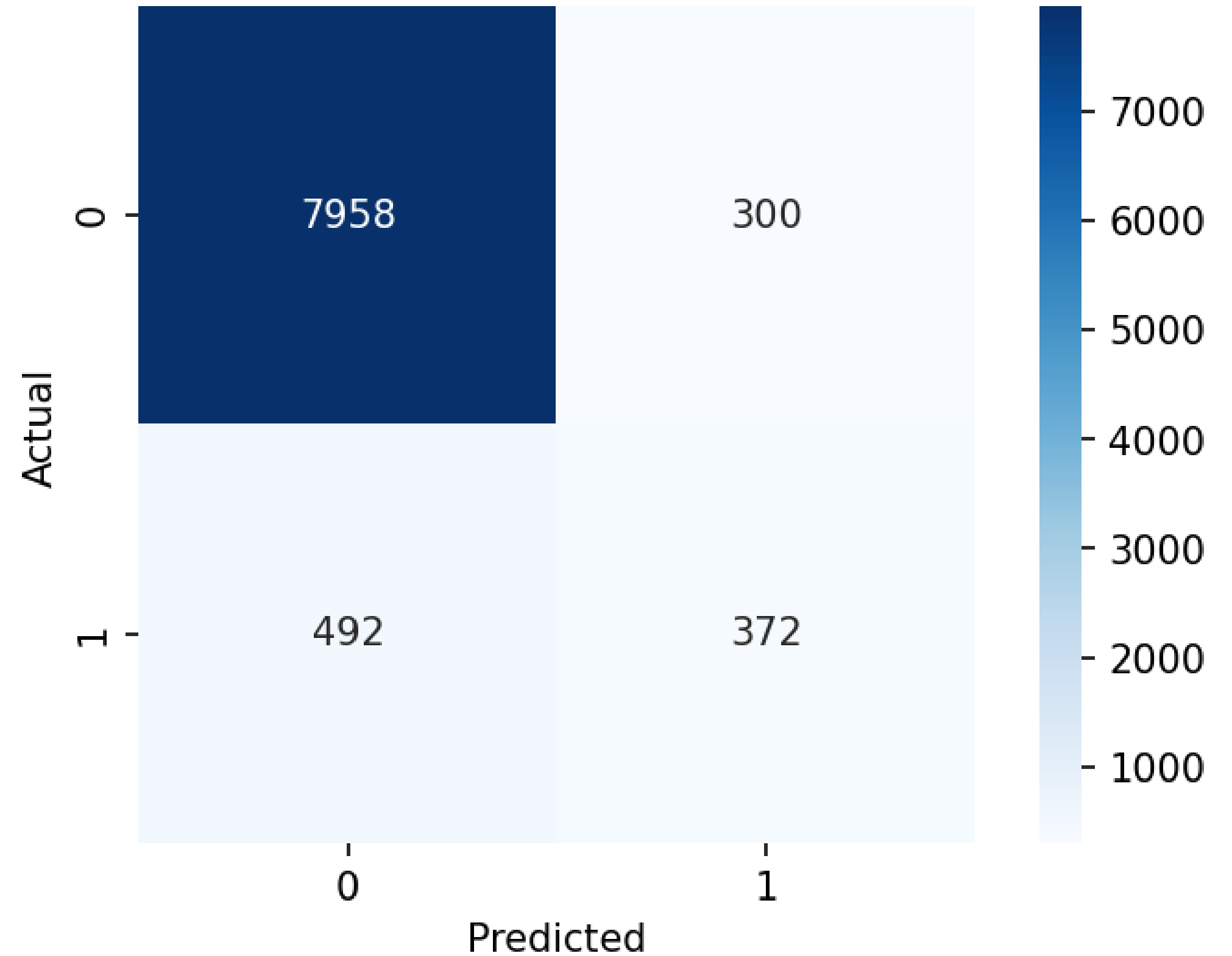
Train

91%

Test

87%

SVM with kernal RBF & C = 100 confusion matrix



DECISION TREE

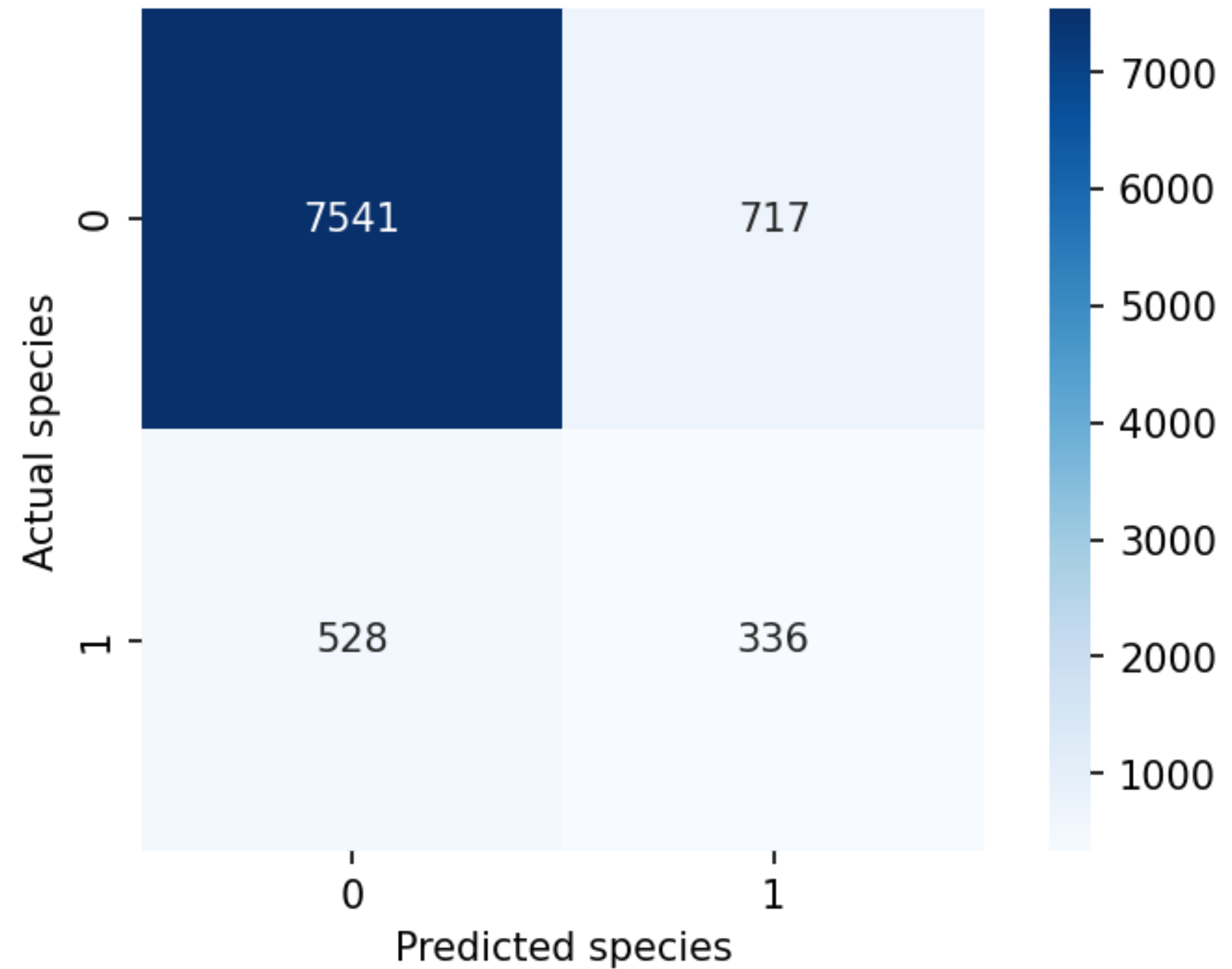
Train

99%

Test

86%

Decision Tree confusion matrix



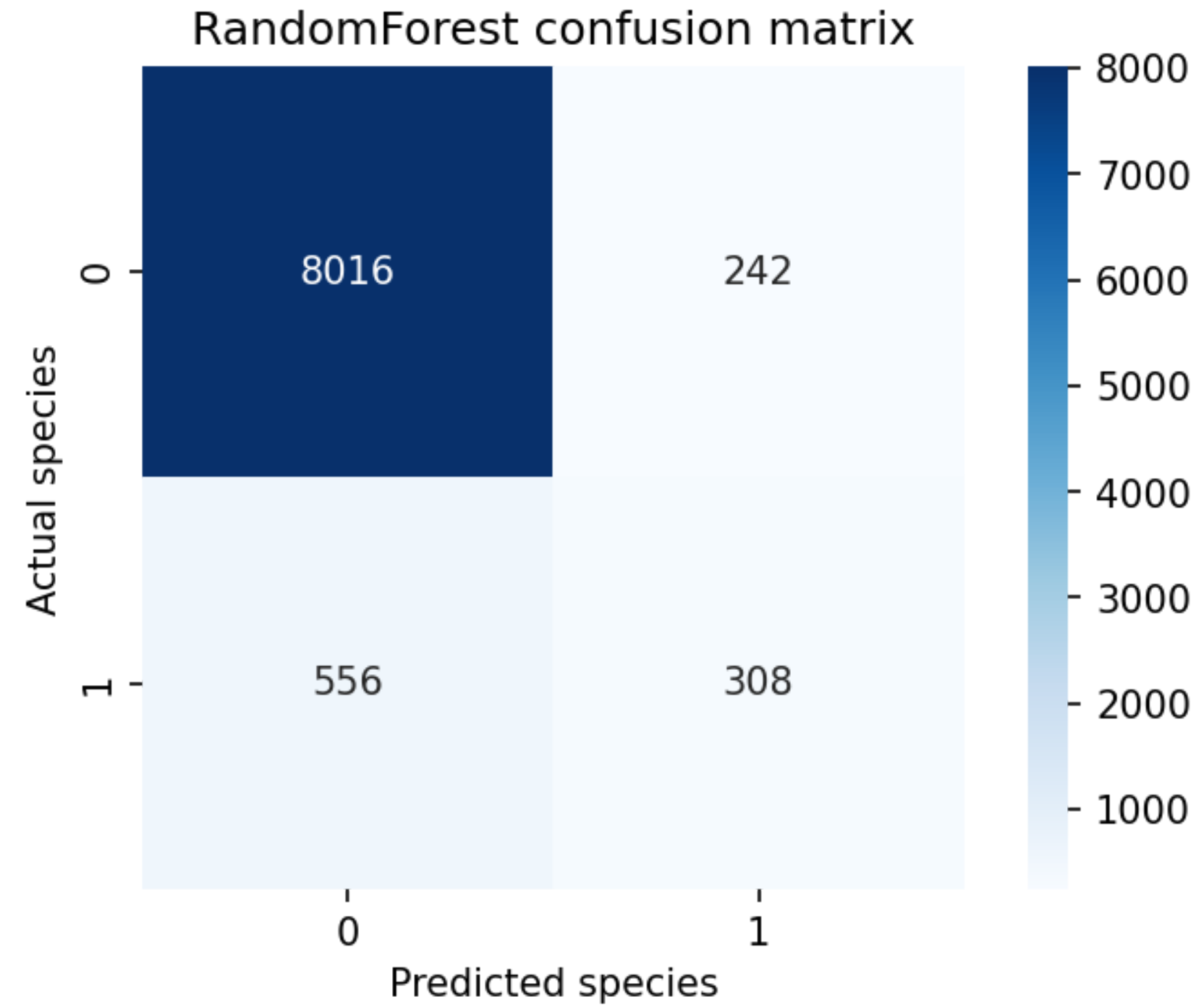
RANDOM FOREST

Train

99%

Test

91%



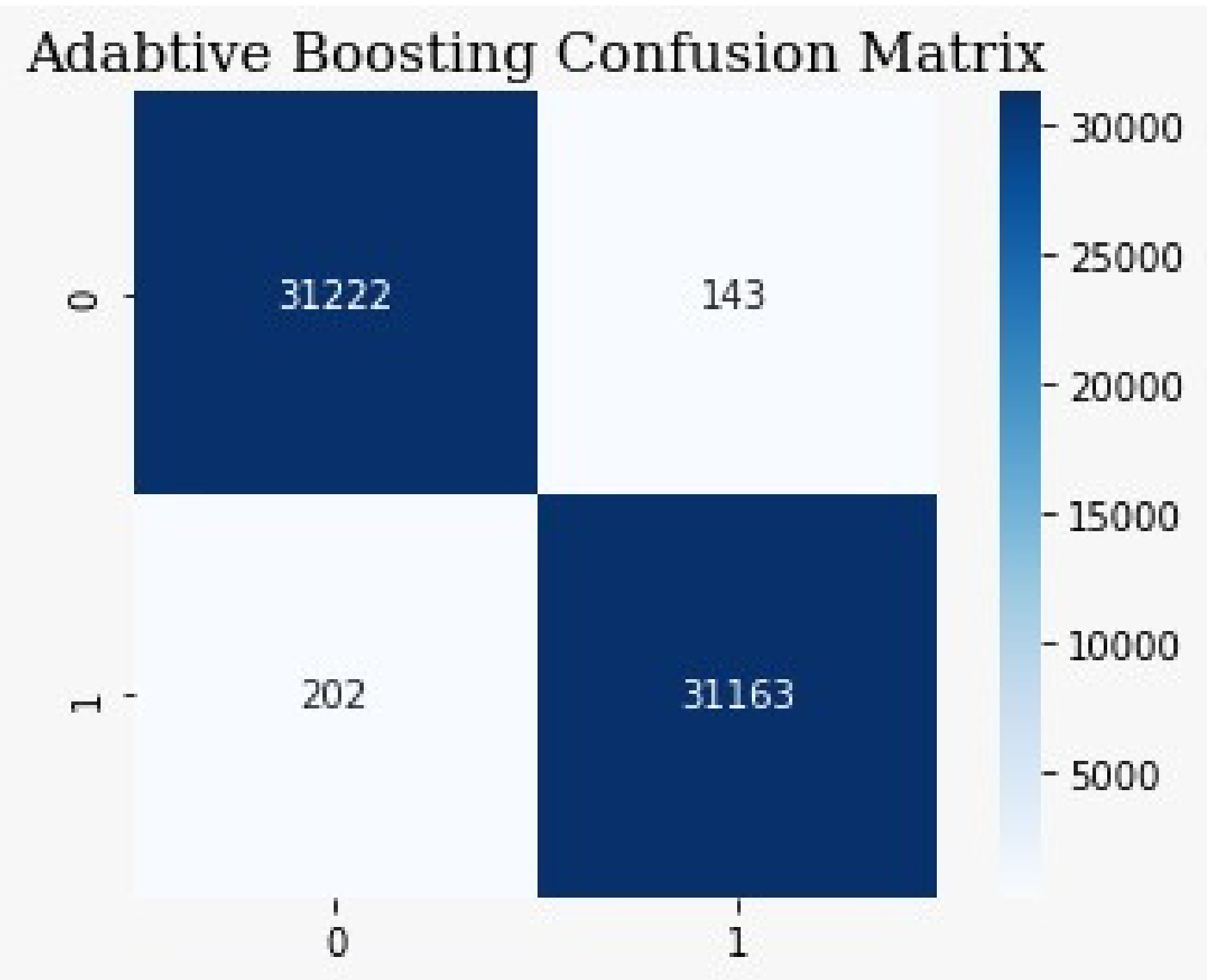
ADABTIVE BOOSTING

Train

99%

Test

89%



GRADIENT BOOSTING

Train

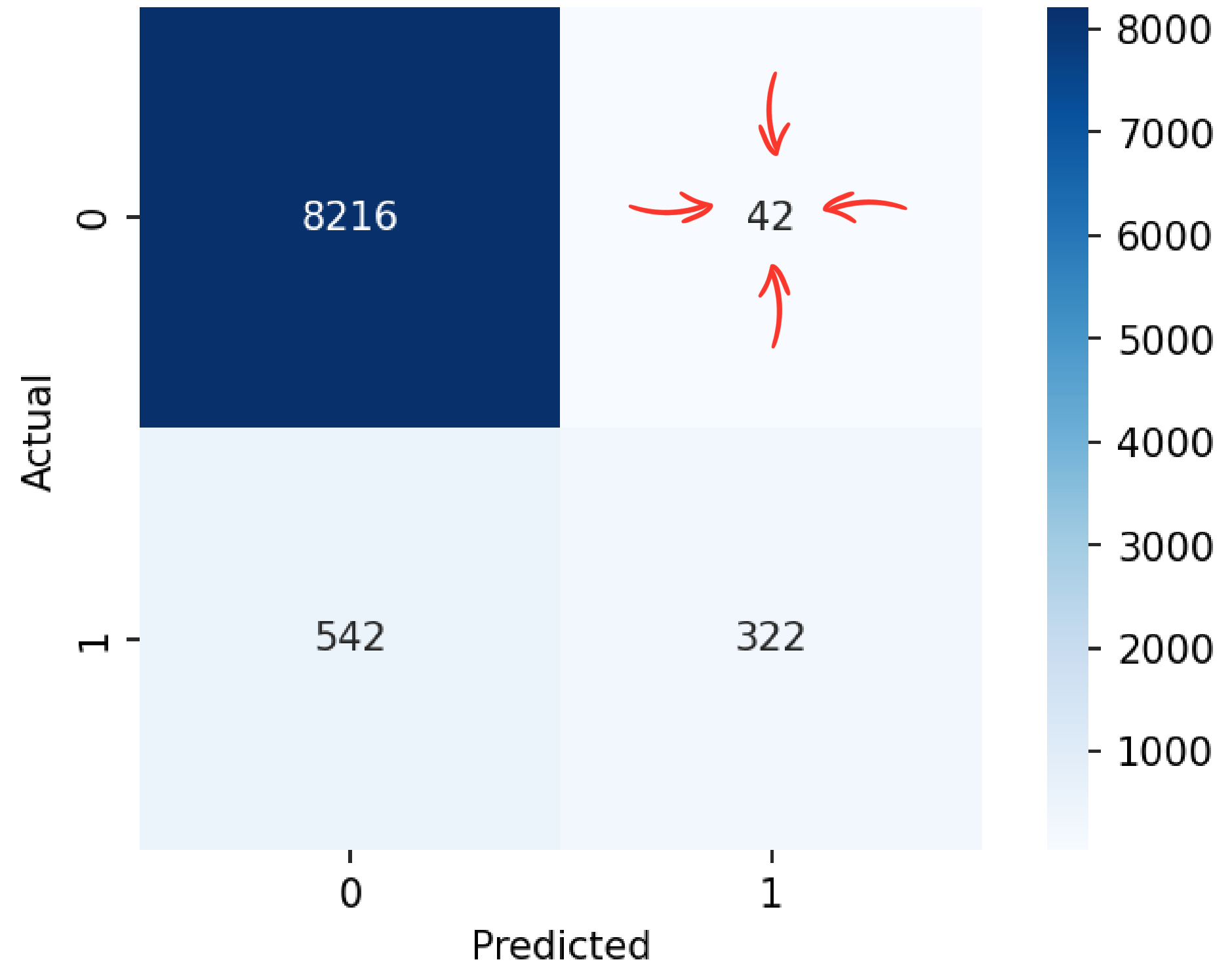
96%

Test


93%



Gradient Boosting confusion matrix



CLASSFICATION REPORT OF GRADIENT BOOSTING

		precision	recall	f1-score	support
	0	0.93	0.99	0.96	7778
	1	0.85	0.36	0.50	888
accuracy				0.93	8666
macro avg		0.89	0.68	0.73	8666
weighted avg		0.92	0.93	0.91	8666

CONCLUSION & Future WORK

- We choose the best performance model that is Gradient Boosting Classifier depends on **accuracy** and **precision, recall**
- We working for the distribution of this model to almost all companies in Saudi Arabia

THANKS !