

# STARTUP FUNDING

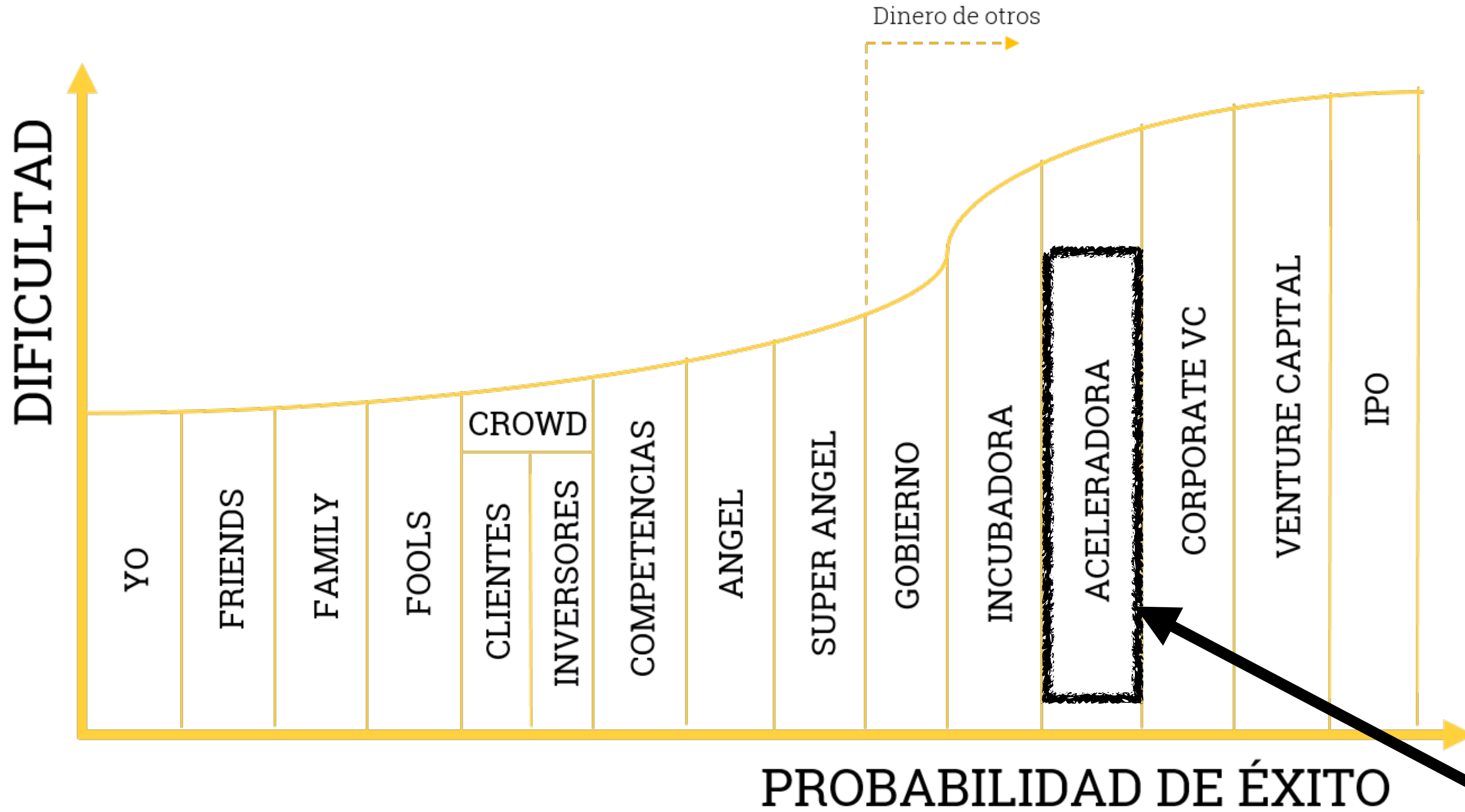
IA PARA PREDECIR EL ÉXITO DE UNA STARTUP  
DAVID GUIX



**LET'S CATCH UP**



¿En qué  
startup  
debemos  
invertir?



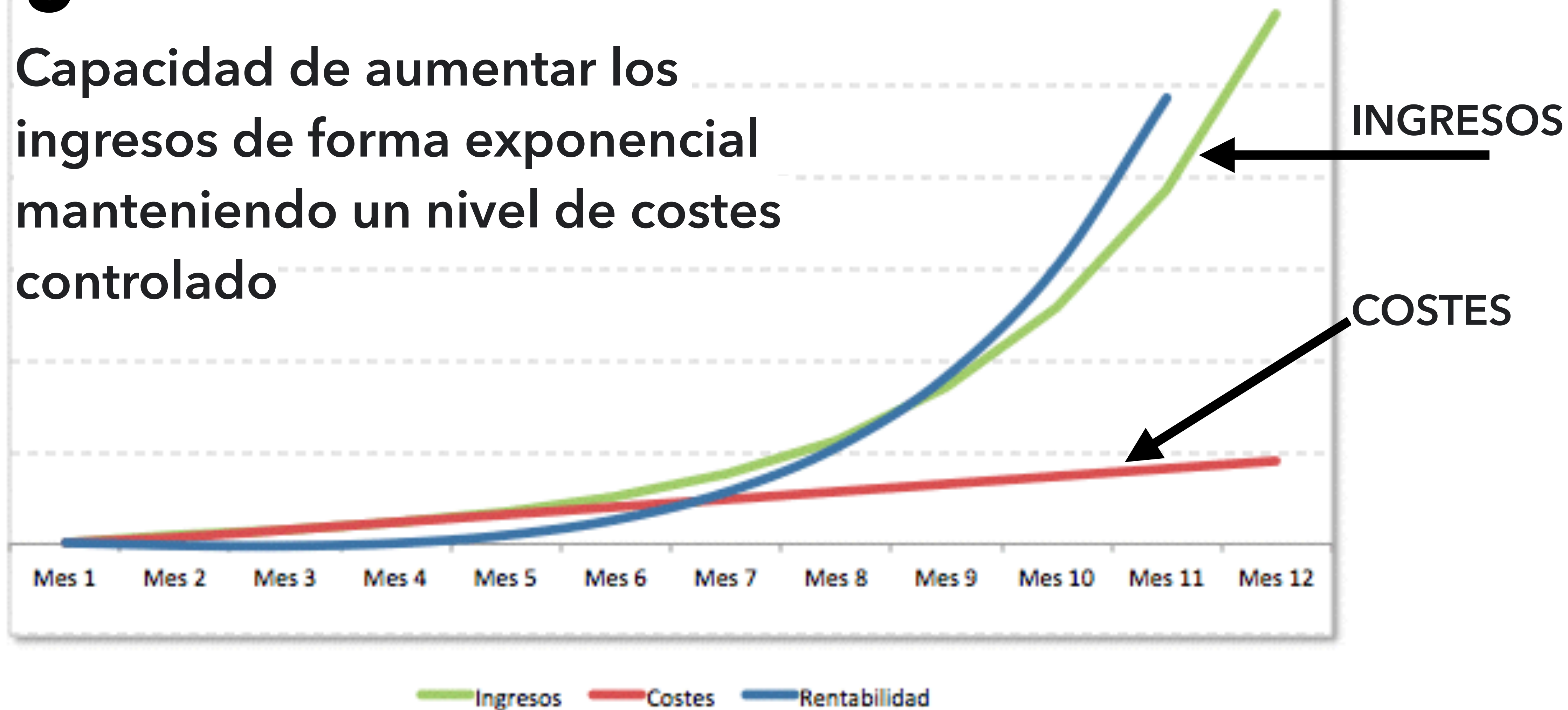
**NEXT LEVEL**  
STARTUP ACCELERATOR

---

***OBJETIVO PRINCIPAL DE UNA  
STARTUP: ESCALAR***

# ¿QUE ES LA ESCALABILIDAD?

Capacidad de aumentar los ingresos de forma exponencial manteniendo un nivel de costes controlado



**¿CÓMO ESCALAR?**

---

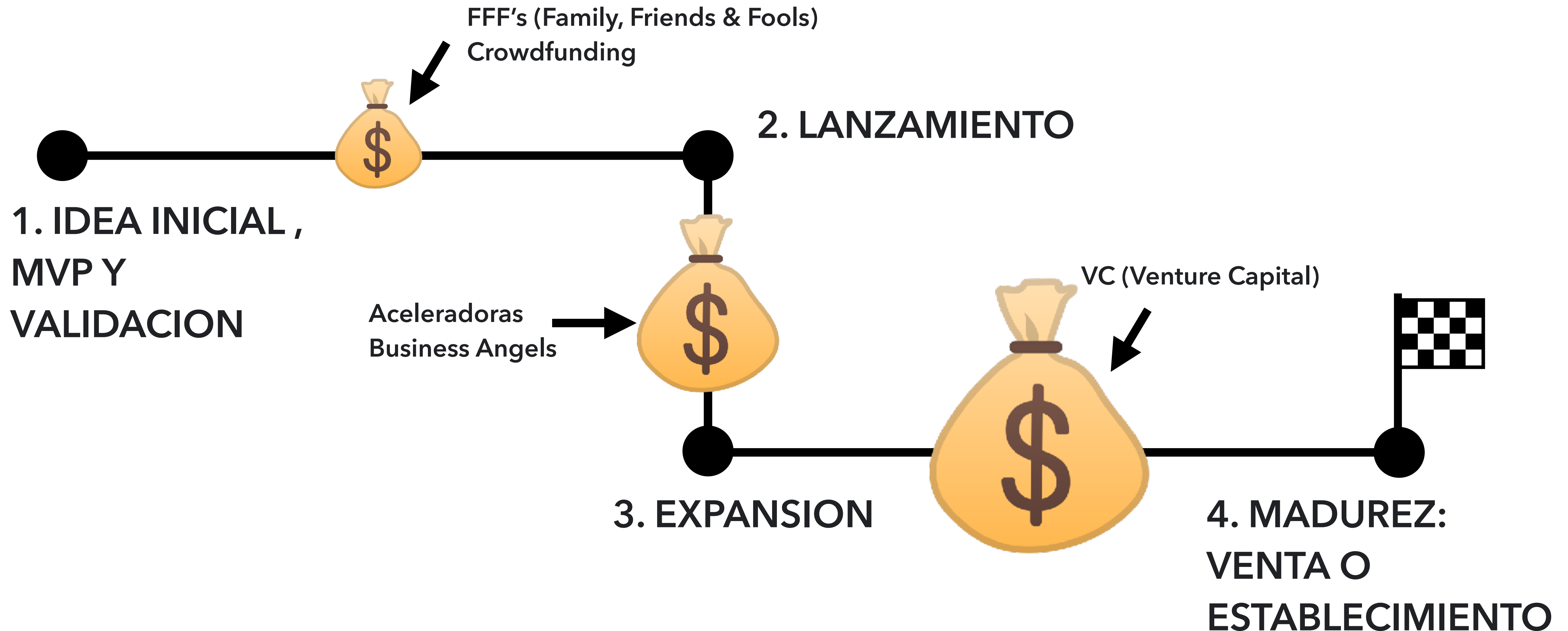
***FUNDING***

***FINANCIACION DE GRUPOS  
INVERSORES***





# CRECIMIENTO - ROAD MAP



# EVALUACION E INVERSION

	AÑO	INVERSION USD	PARTICIPACION %	EVALUACION USD	
FASE 1	2010	80K	5 %	5.4 m	← 2010 VS 2020
FASE 2	2011	3 m	5 %	60 m	
FASE 3	2011	17 m	5 %	346 m	
...	...	...	...	...	
ULTIMA FASE	2020	3.6 b	5 %	76 b	



# **EXPLORATORY DATA ANALYSIS**

# DATOS

- Dataset de Crunchbase 2015
- 54k startups de todo el mundo
- Datos sobre el origen, sector y financiación por startup
- 39 Variables numéricas, fechas y categóricas

# QUE APRENDIMOS CON EL EDA

- Qué países tienen startups con mayor financiación
- Cuales son los sectores más rentables
- En qué momento invertir

**NUEVO OBJETIVO**

# PREDECIR EL ÉXITO DE LAS STARTUPS

- Crear un modelo que prediga la inversión futura
- El modelo debe recomendar invertir o no en las nuevas startups que pidan financiación a la aceleradora



# TENEMOS 100K € Y 3 STARTUPS

	KNEE CREATIONS	ATTUNE FOODS	SKULLY HELMETS
ORIGEN			
QUE VENDEN	Técnica patentada para cirugía de rodilla	Comida orgánica	Cascos de moto de diseño
SECTOR	Health & Wellness	Health & Wellness	Automotive
RONDAS FINANCIACION	3	1	1
PREDICCION	???	???	???



**PROBLEMAS**

# PROBLEMAS

- Los datos que tenemos nos aportan poca info
- No tenemos series temporales de las fases iniciales de las startups mas maduras
- Target poco balanceado y con muchos outliers
- Los algoritmos hacen predicciones poco exactas

**SOLUCIONES**

# SOLUCIONES

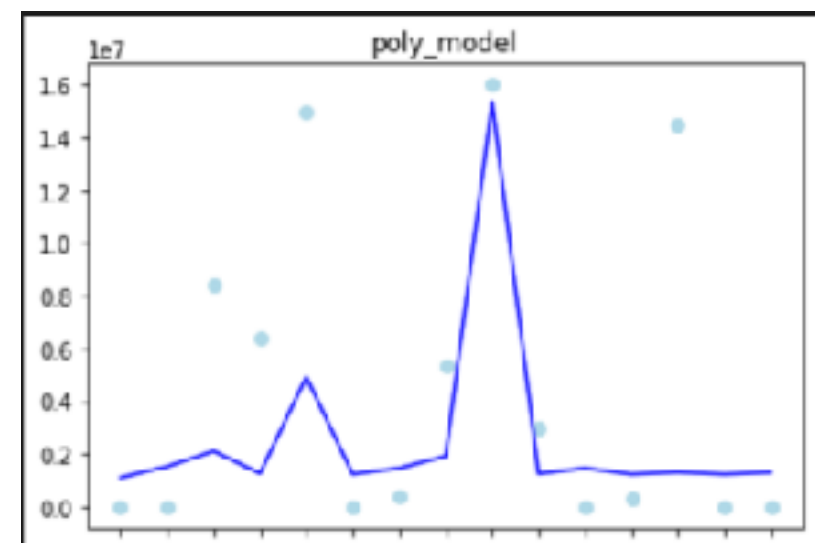
- Sacamos de internet nuevos datos de inversión
- Creamos nuevas features numéricas a partir de categóricas y de fechas
- Filtrado de features por importancia con LR, RF y SKBest
- Eliminamos outliers y balanceamos el target con RUSampler
- Probamos varios modelos de regresión y clasificación
- Generamos un target de clasificación de 3 clases en base al retorno de la inversión: 0 (low), 1 (average), 2 (high)

# LOS MODELOS

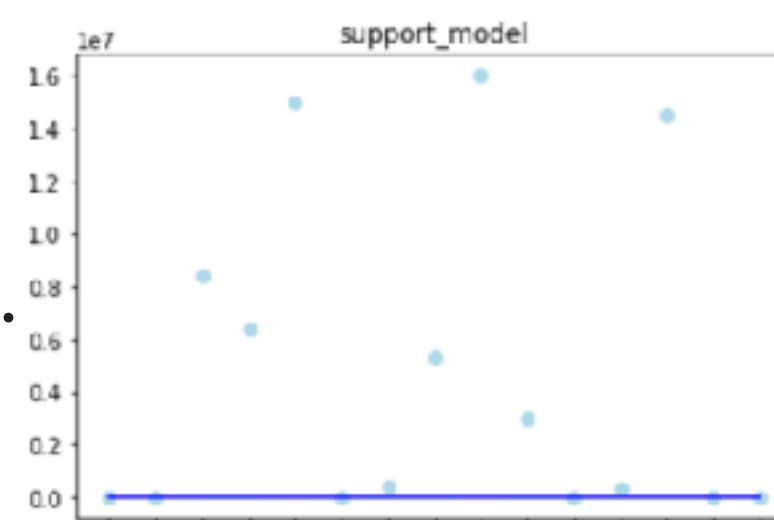
# REGRESION

- Buscamos predecir el total de inversión que recibirán
- Probamos varios algoritmos, RandomForest tiene el menor error pero una Red Neuronal generaliza mejor en las predicciones

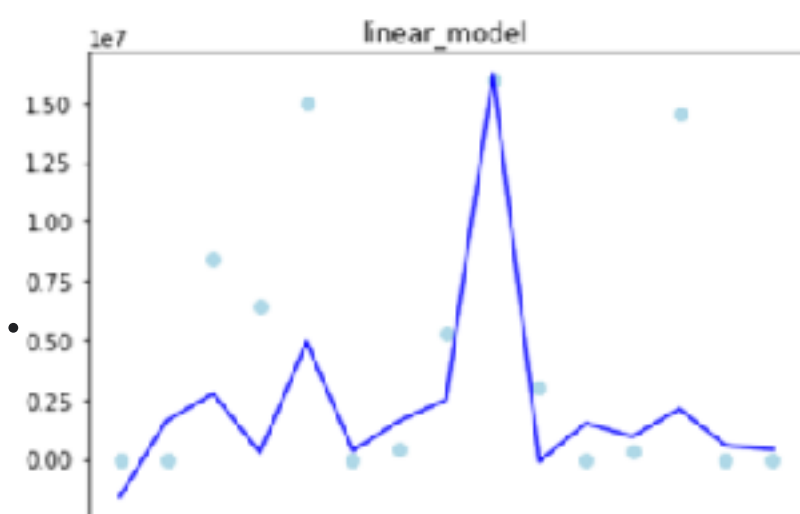
Poly.R



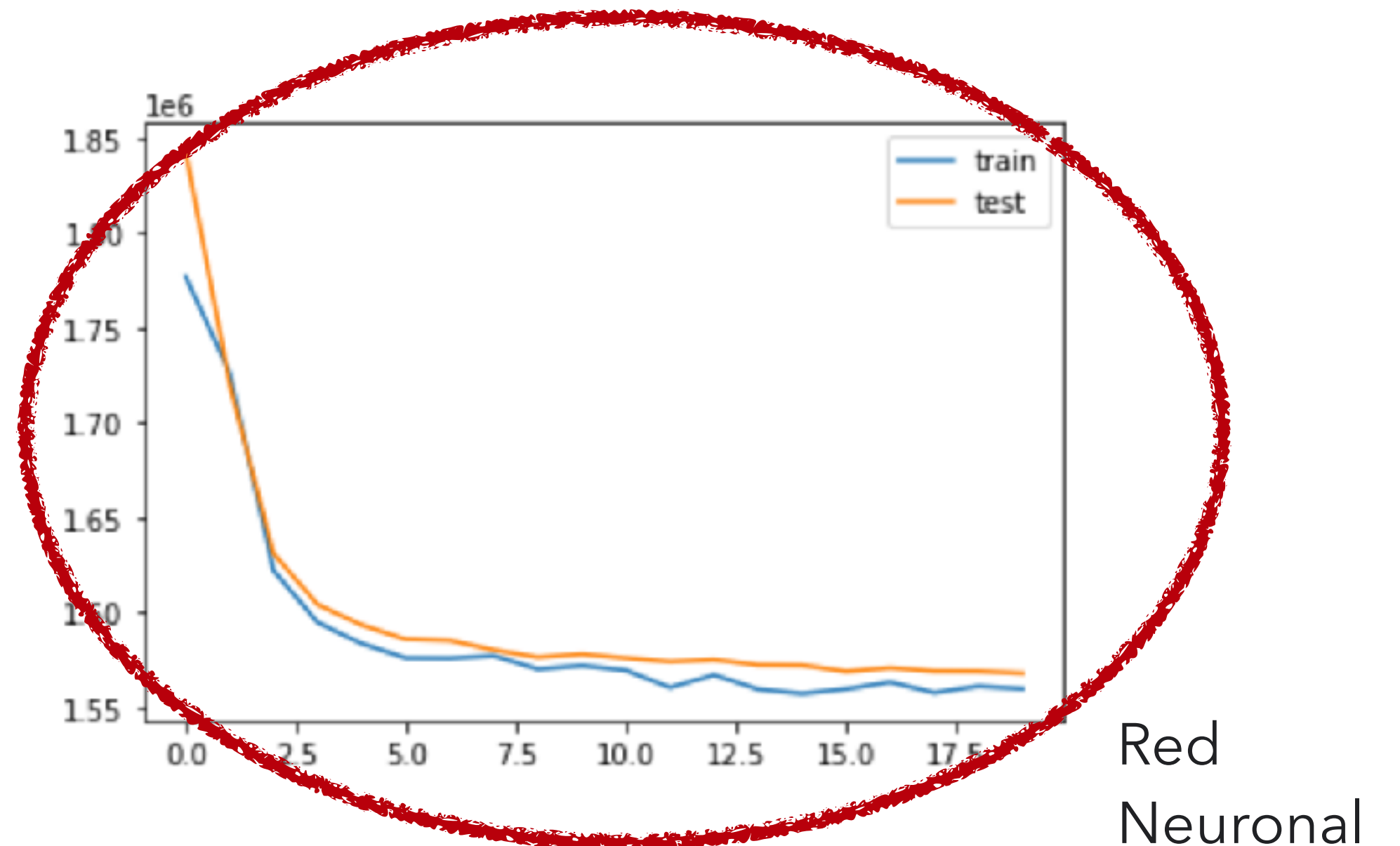
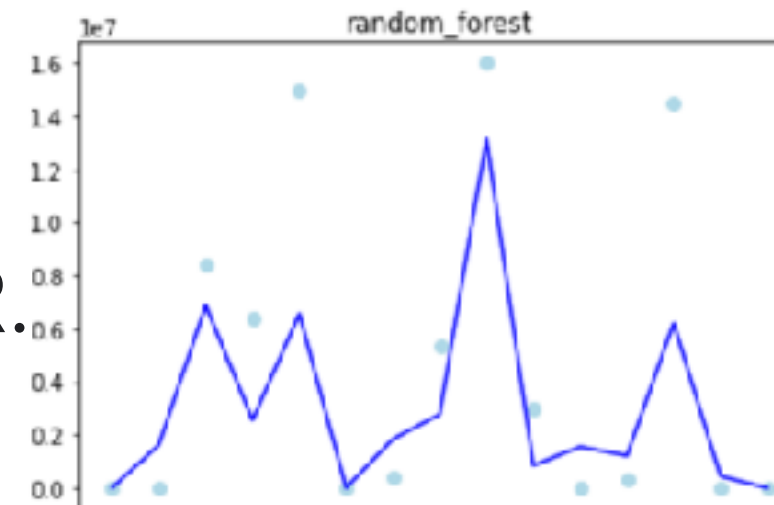
SVR.



LinearR.



RandomFR.



# CLASIFICACION

- Buscamos predecir el grado de inversión: low, average o high
- Probamos varios algoritmos, XGBoost tiene el mejor accuracy pero otra vez la Red Neuronal obtiene mejores predicciones generalizando mucho mejor

XGBoost	0.727775
Ada Boost	0.721414
Neural Network	0.719592
Random Forest	0.711652
Bagging DT	0.705938
KNeighbors	0.675119
GradientBoosting	0.632013



Red  
Neuronal





# SHARKS & DRAGONS CLASSIFIER





# SHARKS & DRAGONS CLASSIFIER

- Añadimos un tercer modelo de clasificación binaria (deal or no deal)
- Scrapeamos datos de internet con el histórico de mas de 1350 startups que se han presentado a los programas de TV *"Sharks Tank"* y *"Dragon's Den"* (USA y UK) desde 2005
- En estos programas un grupo de inversores decide invertir o no en diferentes startups (EXACTAMENTE LO QUE QUEREMOS SABER)
- Han invertido casi 200 m USD con una media de 286k por deal
- Utilizamos una red neuronal con una salida para predecir si la startup tendría éxito en estos programas de TV

**EL ALGORITMO UNICO**

# EL ALGORITMO UNICO



- Hemos creado un algoritmo para unirlos a todos. Su fórmula:

$$y = a W + b W1 + c W2 + d W3 + e W4$$

- Donde a,b y c son los resultados de la red neuronal de clasificacion.  
d= es la predicción de regresión y e= es la predicción del algoritmo Sharks&Dragons
- W son los pesos o coeficientes de cada variable
- El output es un scoring de recomendación de 0 a 5 estrellas



**VOLVEMOS A LAS 3 STARTUPS**

# TENEMOS 100K € Y 3 STARTUPS

	KNEE CREATIONS	ATTUNE FOODS	SKULLY HELMETS
ORIGEN			
QUE VENDEN	Técnica patentada para cirugía de rodilla	Comida orgánica	Cascos de moto de diseño
SECTOR	Health & Wellness	Health & Wellness	Automotive
RONDAS FINANCIACION	3	1	1
PREDICCION	???	???	???

# KNEE CREATIONS

	KNEE CREATIONS
ORIGEN	
QUE VENDEN	Técnica patentada para cirugía de rodilla
SECTOR	Health & Wellness
RONDAS FINANCIACION	3
RECOMENDACION	

	PREDICCION	PESOS
a. Class High	0,793	2
b. Class Average	0,173	1
c. Class Low	0,033	0,1
d. Regresión	4.5 m USD	0,0000005
e. Sharks & Dragons	0,397	1

# ATTUNE FOODS

	ATTUNE FOODS
ORIGEN	
QUE VENDEN	Comida orgánica
SECTOR	Health & Wellness
RONDAS FINANCIACION	1
RECOMENDACION	

	PREDICCION	PESOS
a. Class High	0,339	2
b. Class Average	0,436	1
c. Class Low	0,225	0,1
d. Regresión	478k USD	0,0000005
e. Sharks & Dragons	0,692	1

# SKULLY HELMETS

	SKULLY HELMETS
ORIGEN	
QUE VENDEN	Cascos de moto de diseño
SECTOR	Automotive
RONDAS FINANCIACION	1
RECOMENDACION	

	PREDICCION	PESOS
a. Class High	0,001	2
b. Class Average	0,001	1
c. Class Low	0,993	0,1
d. Regresión	0 USD	0,0000005
e. Sharks & Dragons	0,552	1



# CONCLUSIONES

# TENEMOS 100K € Y 3 STARTUPS

	KNEE CREATIONS	ATTUNE FOODS	SKULLY HELMETS
ORIGEN			
QUE VENDEN	Técnica patentada para cirugía de rodilla	Comida orgánica	Cascos de moto de diseño
SECTOR	Health & Wellness	Health & Wellness	Automotive
RONDAS FINANCIACION	3	1	1
PREDICCION	★★★★☆	★★★★☆	☆☆☆☆

# RESULTADOS PARA 20 STARTUPS

## CLASIFICACION CLASES:

V = 15; X = 3; – = 2

Low Interest	Average Interest	High Interest	actual_venture	
0.999708	0.000188	0.000105	0.0	V
0.999447	0.000188	0.000365	0.0	V
0.033373	0.173600	0.793027	6500000.0	V
0.225498	0.435556	0.338946	5000000.0	–
0.707494	0.177856	0.114651	64795.0	–
0.944985	0.019990	0.035025	0.0	V
0.999934	0.000041	0.000025	0.0	V
0.993707	0.005218	0.001075	0.0	V
0.993696	0.003443	0.002861	0.0	V
0.654208	0.164437	0.181355	2000000.0	X
0.998356	0.000542	0.001102	0.0	V
0.607553	0.248085	0.144362	0.0	V
0.111985	0.182105	0.705910	4800000.0	V
0.014626	0.248612	0.736762	13345667.0	V
0.811774	0.085667	0.102559	0.0	V
0.999849	0.000036	0.000116	0.0	V
0.822062	0.124488	0.053450	715000.0	X
0.315941	0.266802	0.417258	0.0	X
0.705801	0.201449	0.092750	0.0	V
0.959204	0.017912	0.022884	0.0	V

## REGRESION:

V = 14; X = 5; – = 1

predicted_NeuralNetwork	actual_venture	
1.972944e-01	0.0	V
2.923777e-01	0.0	V
4.558084e+06	6500000.0	V
4.780959e+05	5000000.0	–
2.359650e-01	64795.0	X
3.672479e+05	0.0	X
2.254927e-01	0.0	V
2.563784e-01	0.0	V
2.202637e-01	0.0	V
2.148191e-01	2000000.0	X
1.803815e-01	0.0	V
1.006306e+05	0.0	X
3.419121e+06	4800000.0	V
5.514424e+06	13345667.0	V
2.266819e-01	0.0	V
2.884638e-01	0.0	V
2.146751e-01	715000.0	X
1.948969e-01	0.0	V
1.924078e-01	0.0	V
2.570059e-01	0.0	V

## CLAS BINARIA SHARKS & DRAGONS:

V = 5; X = 11; – = 4

Sharks_and_Dragons	actual_venture	
0.518695	0.0	–
0.535030	0.0	–
0.397517	6500000.0	X
0.692109	5000000.0	V
0.631560	64795.0	V
0.601084	0.0	X
0.618361	0.0	X
0.625371	0.0	X
0.545289	0.0	–
0.371615	2000000.0	X
0.552096	0.0	X
0.595465	0.0	X
0.591428	4800000.0	V
0.443691	13345667.0	X
0.487639	0.0	V
0.600604	0.0	X
0.565286	715000.0	V
0.467874	0.0	–
0.635233	0.0	X
0.611452	0.0	X



# RESULTADOS PARA 20 STARTUPS

#df_new_pred[['Low Interest', 'Average Interest', 'High Interest']] = predict_RNC											
df_new_pred['Sharks_and_Dragons'] = y_pred_SaD											
'''											
startup1_SaD		from_founding_to_last_funding	founded_year	funding_rounds	seed	undisclosed	from_founding_to_funding	from_first_to_last_funding	math_expectation_country	product_crowdfunding	
startup2_SaD	0	6.341760e+16	2010.0	1.0	325000.0	0.0	6.341760e+16	0.000000e+00	77.650555	0.0	
startup3_SaD	1	3.153800e+16	2006.0	1.0	1107075.0	0.0	3.153800e+16	0.000000e+00	0.000000	0.0	
	2	1.563840e+17	2007.0	3.0	0.0	0.0	7.084800e+16	8.553600e+16	119.305251	0.0	
print(mse_te	3	1.252000e+17	2006.0	1.0	0.0	0.0	1.252000e+17	0.000000e+00	119.305251	0.0	
	4	7.776000e+15	2005.0	1.0	0.0	0.0	7.776000e+15	0.000000e+00	12.644516	0.0	
	5	4.224960e+16	2008.0	2.0	140000.0	0.0	0.000000e+00	4.224960e+16	0.000000	0.0	
	6	8.261760e+16	2012.0	1.0	400000.0	0.0	8.261760e+16	0.000000e+00	32.385041	0.0	
pred_startup	7	1.720000e+15	2012.0	1.0	200000.0	0.0	1.720000e+15	0.000000e+00	119.305251	0.0	
pred_startup	8	0.000000e+00	2006.0	1.0	163687.0	0.0	0.000000e+00	0.000000e+00	79.546391	0.0	
pred_startup	9	4.805120e+16	2013.0	1.0	0.0	0.0	4.805120e+16	0.000000e+00	40.831250	0.0	
'''	10	4.803840e+16	2013.0	1.0	0.0	0.0	4.803840e+16	0.000000e+00	119.305251	1682454.0	
	11	2.946240e+16	2009.0	1.0	0.0	0.0	2.946240e+16	0.000000e+00	119.305251	0.0	
	12	7.032960e+16	2008.0	2.0	0.0	0.0	2.678400e+15	6.765120e+16	119.305251	0.0	
post_ipo_debt	debt_financing	predicted_NeuralNetwork	actual_venture	diff_pred_actual_NN	Low Interest	Average Interest	High Interest	Sharks_and_Dragons	Algoritmo_Unico	119.305251	0.0
										77.650555	0.0
0.0	0.0	1.972944e-01	0.0	1.972944e-01	0.999708	0.000188	0.000105	0.518695	0.62	119.305251	0.0
0.0	0.0	2.923777e-01	0.0	2.923777e-01	0.999447	0.000188	0.000365	0.535030	0.64	119.305251	0.0
0.0	0.0	4.558084e+06	6500000.0	-1.941916e+06	0.033373	0.173600	0.793027	0.397517	4.44	79.546391	0.0
0.0	0.0	4.780959e+05	5000000.0	-4.521904e+05	0.225498	0.435556	0.338946	0.692109	2.07	119.305251	0.0
0.0	0.0	2.358650e-01	64795.0	-6.479476e+04	0.707494	0.177856	0.114651	0.631560	1.11	77.650555	0.0
0.0	0.0	3.672479e+05	0.0	3.672479e+05	0.944935	0.019990	0.035025	0.601084	0.97		
0.0	0.0	2.254927e-01	0.0	2.254927e-01	0.999934	0.000041	0.000025	0.618361	0.72		
0.0	0.0	2.563784e-01	0.0	2.563784e-01	0.993707	0.005218	0.001075	0.625371	0.73		
0.0	0.0	2.202637e-01	0.0	2.202637e-01	0.993696	0.003443	0.002861	0.545289	0.65		
0.0	0.0	2.148191e-01	2000000.0	-2.000000e+06	0.654208	0.164437	0.181355	0.371615	0.96		
0.0	0.0	1.803815e-01	0.0	1.803815e-01	0.998356	0.000542	0.001102	0.552096	0.65		
0.0	0.0	1.006306e+05	0.0	1.006306e+05	0.607553	0.248085	0.144362	0.595465	1.24		
0.0	0.0	3.419121e+06	4800000.0	-1.380879e+06	0.111935	0.182105	0.705910	0.591428	3.91		
0.0	1686193.0	5.514424e+06	13345667.0	-7.831243e+06	0.014626	0.248612	0.736762	0.443691	4.92		
0.0	0.0	2.266819e-01	0.0	2.266819e-01	0.811774	0.085667	0.102559	0.487639	0.86		
0.0	0.0	2.884638e-01	0.0	2.884638e-01	0.999849	0.000036	0.000116	0.600604	0.70		
0.0	0.0	2.146751e-01	715000.0	-7.149996e+05	0.822062	0.121488	0.053450	0.565286	0.88		
0.0	0.0	1.948969e-01	0.0	1.948969e-01	0.315941	0.266802	0.417258	0.467874	1.60		
0.0	0.0	1.924078e-01	0.0	1.924078e-01	0.705801	0.201449	0.092750	0.635233	1.09		
0.0	0.0	2.570059e-01	0.0	2.570059e-01	0.959204	0.017912	0.022884	0.611452	0.77		

Algoritmo_Unico	actual_venture
0.62	0.0
0.64	0.0
4.44	6500000.0
2.07	5000000.0
1.11	64795.0
0.97	0.0
0.72	0.0
0.73	0.0
0.65	0.0
0.96	2000000.0
0.65	0.0
1.24	0.0
3.91	4800000.0
4.92	13345667.0
0.86	0.0
0.70	0.0
0.88	715000.0
1.60	0.0
1.09	0.0
0.77	0.0

✓

✓

✓

—

—

✓

✓

✓

✓

✗

✓

✓

✓

✓

✓

✓

✗

—

✓

✓

ALGORITMO UNICO:

V = 15;

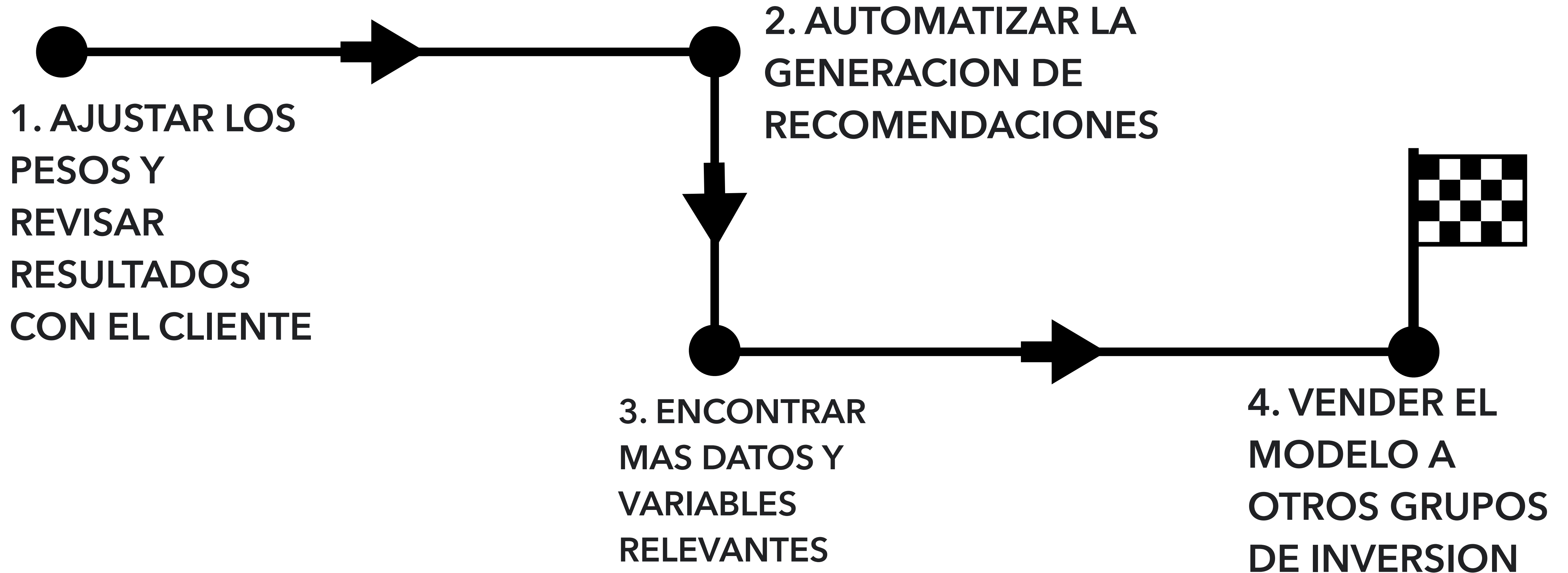
X = 2;

— = 3

# NEXT STEPS

# NEXT STEPS

## ROAD MAP



**GRACIAS**