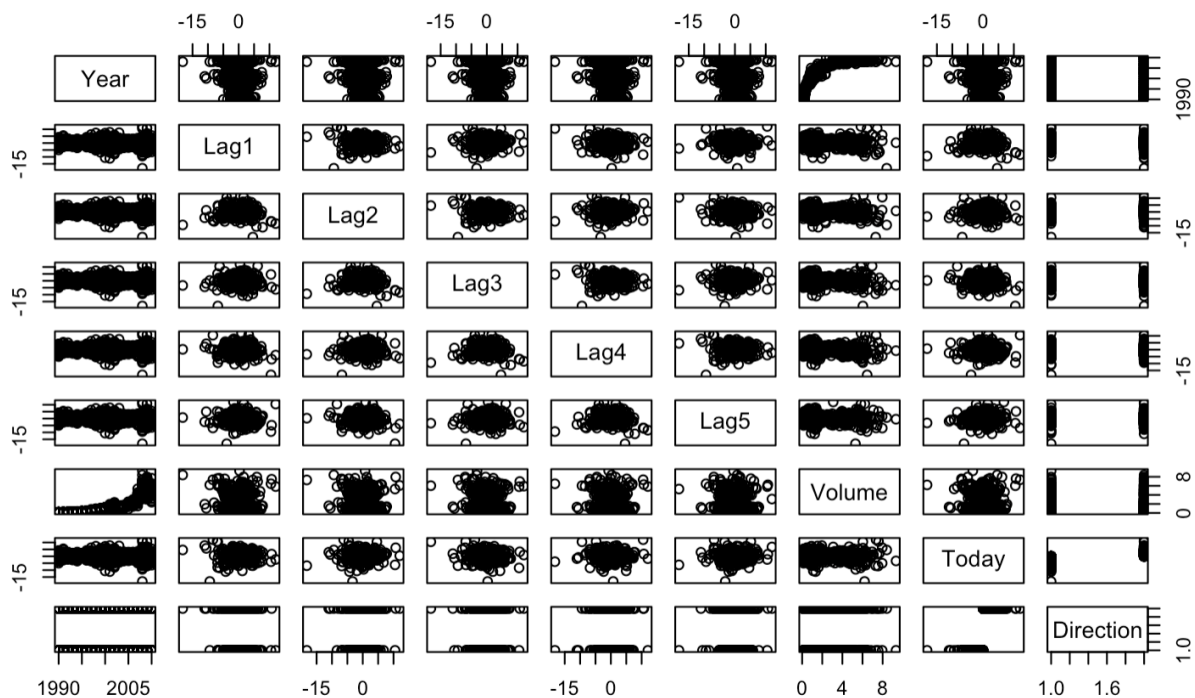


Jorge J. Sosa Briseño

A01749489

October 19, 2023

Regresión Logística



```

#Cálculo del modelo logístico
##Modelo con todos los predictores, excluyendo
##"Today"
modelo.log.m <- glm(Direction ~ . -Today, data
+ Weekly, family = binomial)
summary(modelo.log.m)
contrasts(Direction)
confint(object = modelo.log.m, level = 0.95)

```

```

Call:
glm(formula = Direction ~ . - Today, family = binomial, data = Weekly)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.7071  -1.2578   0.9941   1.0873   1.4665

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) 17.225822  37.890522   0.455  0.6494
Year         -0.008500   0.018991  -0.448  0.6545
Lag1         -0.040688   0.026447  -1.538  0.1239
Lag2          0.059449   0.026970   2.204  0.0275 *
Lag3         -0.015478   0.026703  -0.580  0.5622
Lag4         -0.027316   0.026485  -1.031  0.3024
Lag5         -0.014022   0.026409  -0.531  0.5955
Volume        0.003256   0.068836   0.047  0.9623
---

```

El modelo indica que solo "Lag2" tiene un efecto significativo en la predicción de la dirección del mercado, mientras que las demás variables no aportan información significativa. Sin embargo, el modelo en su conjunto puede no ser suficientemente bueno para predecir con precisión la dirección del mercado, y podrían ser necesarios enfoques más sofisticados o características adicionales para mejorar la capacidad de predicción.

```

integer(0)

Call:
glm(formula = Direction ~ Lag2, family = binomial, data = Weekly,
    subset = datos.entrenamiento)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.536  -1.264   1.021   1.091   1.368

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  0.20326   0.06428   3.162  0.00157 **
Lag2         0.05810   0.02870   2.024  0.04298 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

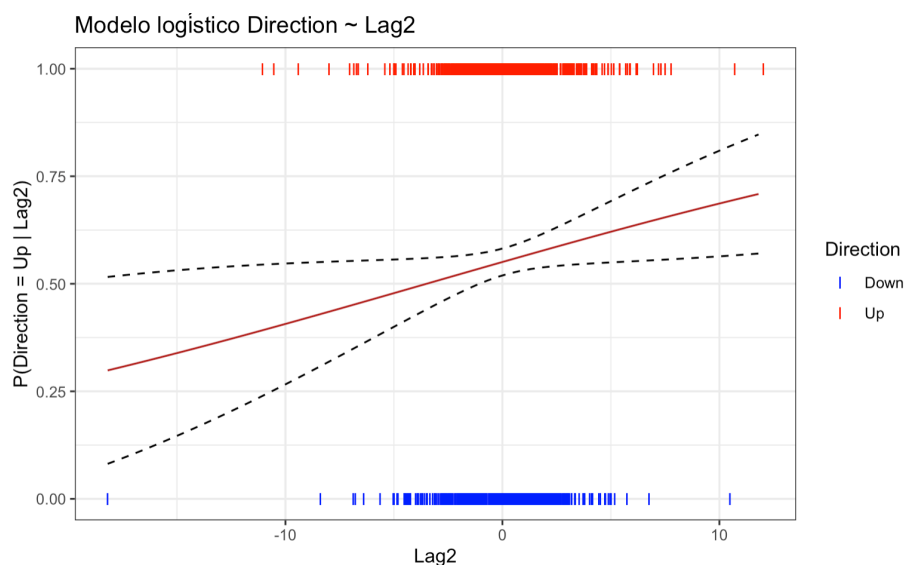
(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 1354.7  on 984  degrees of freedom
Residual deviance: 1350.5  on 983  degrees of freedom
AIC: 1354.5

Number of Fisher Scoring iterations: 4

```

Este modelo logístico utiliza solo la variable "Lag2" como predictor y sugiere que un aumento en "Lag2" está asociado con un aumento en la probabilidad de que la dirección del mercado sea "Up". El modelo tiene un ajuste razonable a los datos, pero su capacidad predictiva puede mejorarse al considerar más variables o métodos más avanzados. Como se muestra a continuación con sus respectivos intervalos de confianza:



```

{r}
# Chi cuadrada: Se evalúa la significancia del modelo con predictores con respecto al modelo nulo ("Residual deviance" vs
"Null deviance"). Si valor p es menor que alfa será significativo.
anova(modelo.log.s, test = 'Chisq')

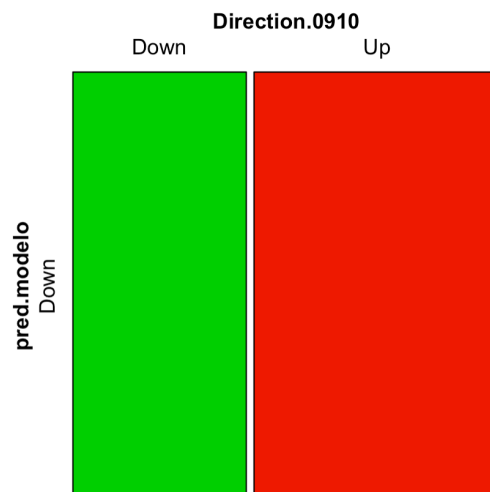
```

	Df	Deviance	Resid.	Df	Resid.	Dev	Pr(>Chi)
NULL			984		1354.7		
Lag2	1	4.1666	983		1350.5	0.04123	*

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

- El modelo nulo (sin predictores) tiene una devianza nula (Null deviance) de 1354.7 y se calcula sobre 984 grados de libertad.
- El modelo que incluye la variable "Lag2" tiene una devianza residual (Residual deviance) de 1350.5 y se calcula sobre 983 grados de libertad.
- La columna "Pr(>Chi)" muestra el valor p de la prueba de chi-cuadrada. En este caso, el valor p es 0.04123, que es menor que el nivel de significancia alfa (generalmente 0.05), lo que indica que el modelo con "Lag2" es significativamente mejor que el modelo nulo.
- Además, el asterisco (*) al lado del valor p indica que la diferencia entre el modelo con "Lag2" y el modelo nulo es estadísticamente significativa a un nivel de significancia de 0.05.

La prueba de chi-cuadrada muestra que el modelo que incluye la variable "Lag2" es significativamente mejor que un modelo nulo sin predictores, lo que sugiere que "Lag2" tiene un efecto significativo en la predicción de la dirección del mercado.



- Para la categoría "Down", el modelo ha realizado 43 predicciones correctas de "Down" y 61 predicciones incorrectas de "Up". Esto significa que 43 observaciones reales que fueron "Down" fueron correctamente clasificadas como "Down" por el modelo, pero 61 observaciones reales que fueron "Down" fueron incorrectamente clasificadas como "Up".

- La fila "[1] 0.4134615" parece ser la tasa de precisión general del modelo para el período 2009-2010. En este caso, el valor 0.4134615 indica que el modelo tiene una precisión global del 41.35%. Esto significa que aproximadamente el 41.35% de las observaciones reales fueron clasificadas correctamente por el modelo en el conjunto de prueba.

La matriz de confusión y la tasa de precisión son indicadores importantes para evaluar el rendimiento de un modelo de clasificación.