永豐金證券
SinoPac Securities

千人千面的證券自動化服務

搭上動能列車-即時話題股篩選器

Mentor：

呂正傑 業師

鄭維 助教

指導教授：

交大 戴天時教授

東吳 蔡芸琤教授

組員：

東吳巨資 戴緯綾

東吳巨資 陳首暄

東吳巨資 李重諺

東吳巨資 何佳穎

台大財金 陳若昕

目錄

CONTENTS

01

PART

題目說明

題目說明

現今資訊量龐大，投資者無法一時掌握所有資訊，我們希望能夠以最簡單明瞭的方式來提供給投資者一些選擇，讓投資者能夠用最快速的時間清楚掌握投資的方向。起初我們這組皆對Mini LED相關產業有興趣，因此我們決定以Mini LED 的概念股做為出發點，接著利用機器學習來預測股價走勢，並加入其他相關公司財務資訊，讓投資者可以在一個網站就得知所有資訊，做出最有利的投資。

02
PART

模型成果

專案進度規劃

01 爬蟲

- 公開資訊觀測站-財務報表
- Google NEWS
- 股票線圖

02 機器學習情緒分析

- LSTM
- SVR
- Linear Regression
- Random Forest
- KNN
- Sentence Embedding

03 網頁

- 股票趨勢預測
- K線圖
- 流動資產與負債佔比
- PE Ratio
- 現金殖利率
- 毛利率
- 營業利益率
- 淨利率

參數

	date	high	low	open	volume	adj close	ma5	ma10	ma30	macd	ema	year	month	date_1	day	return	close
1196	2021-05-04	30.500000	26.900000	29.000000	576426814.0	27.400000	30.970000	29.790	24.928333	2.564320	28.557071	2021	5	4	1	-2.450001	27.400000
1197	2021-05-05	28.450001	26.350000	27.400000	366398239.0	27.000000	30.150000	29.800	25.191667	2.215847	27.519024	2021	5	5	2	-0.400000	27.000000
1198	2021-05-06	28.150000	25.049999	27.600000	417870867.0	25.700001	28.620000	29.460	25.411667	1.813871	26.306342	2021	5	6	3	-1.299999	25.700001
1199	2021-05-07	28.200001	25.350000	26.450001	418288185.0	28.150000	27.620000	29.475	25.683333	1.673704	27.535447	2021	5	7	4	2.449999	28.150000
1200	2021-05-10	28.000000	26.650000	27.950001	287212738.0	26.850000	27.020000	29.275	25.920000	1.441109	27.078483	2021	5	10	0	-1.299999	26.850000
1201	2021-05-11	25.400000	24.200001	24.850000	364676340.0	24.200001	26.380000	28.675	26.043333	1.031057	25.159495	2021	5	11	1	-2.650000	24.200001
1202	2021-05-12	24.200001	21.799999	23.850000	588538141.0	22.000000	25.380000	27.765	26.126667	0.522543	23.053165	2021	5	12	2	-2.200001	22.000000
1203	2021-05-13	22.750000	20.100000	20.600000	472602880.0	21.799999	24.600000	26.610	26.181667	0.102225	22.217721	2021	5	13	3	-0.200001	21.799999
1204	2021-05-14	23.700001	21.350000	22.799999	417469845.0	21.850000	23.340000	25.480	26.200000	-0.224260	21.972574	2021	5	14	4	0.050001	21.850000
1205	2021-05-17	21.799999	19.700001	19.799999	377943873.0	20.000000	21.970000	24.495	26.150000	-0.625075	20.657525	2021	5	17	0	-1.850000	20.000000
1206	2021-05-18	22.000000	21.150000	21.450001	314350373.0	22.000000	21.530000	23.955	26.181667	-0.772437	21.552508	2021	5	18	1	2.000000	22.000000
1207	2021-05-19	23.950001	21.850000	22.200001	733202518.0	23.250000	21.780000	23.580	26.255000	-0.779374	22.684169	2021	5	19	2	1.250000	23.250000
1208	2021-05-20	23.150000	21.650000	22.750000	351994907.0	21.850000	21.790000	23.195	26.210000	-0.887608	22.128057	2021	5	20	3	-1.400000	21.850000
1209	2021-05-21	22.799999	21.350000	22.500000	365876551.0	21.700001	21.760000	22.550	26.155000	-0.974257	21.842686	2021	5	21	4	-0.150000	21.700001
1210	2021-05-24	22.900000	20.900000	21.400000	311048122.0	22.600000	22.280000	22.125	26.088333	-0.959247	22.347562	2021	5	24	0	0.900000	22.600000
1211	2021-05-25	23.900000	22.799999	23.150000	447228641.0	22.850000	22.450000	21.990	25.948333	-0.916613	22.682521	2021	5	25	1	0.250000	22.850000
1212	2021-05-26	23.200001	22.200001	22.950001	227252747.0	22.200001	22.240001	22.010	25.835000	-0.924616	22.360841	2021	5	26	2	-0.650000	22.200001
1213	2021-05-27	23.049999	21.750000	21.850000	332208753.0	22.500000	22.370000	22.080	25.758334	-0.896418	22.453614	2021	5	27	3	0.299999	22.500000
1214	2021-05-28	23.700001	23.000000	23.200001	324937621.0	23.500000	22.730000	22.245	25.698334	-0.784337	23.151205	2021	5	28	4	1.000000	23.500000
1215	2021-05-31	25.850000	23.750000	24.250000	738264242.0	25.850000	23.380000	22.830	25.698334	-0.500122	24.950402	2021	5	31	0	2.350000	25.850000

Sentence Embedding

163	2021-05-14	國家隊進場 資金轉進哪類股？萬寶投顧陳子榕指出，台股屋漏偏逢連夜雨，近日本土疫情升溫，興達電...	0.051574	-0.041362	-0.012775	-0.001936	0.041267	-0.066078	0.008211	0.005494	-0.008161	0.026479	0.011382	-0.009064	0.004501	0.006400	-0.012287	0.044338	0.039764	-0.024154	-(
164	2021-05-15	官股單周砸208億元穩盤 順勢調節航運雙雄國內疫情升溫，拖累台股大盤指數跳水，單周重挫逾1...	0.046722	-0.010780	-0.013240	-0.000755	-0.002787	-0.043575	0.045596	0.014485	-0.012939	0.019103	-0.027624	0.000393	0.012572	0.008757	-0.018133	0.033509	0.019178	-0.015481	-(
165	2021-05-17	nan描繪疫後智慧生活新藍圖 友達SID秀尖端顯示技術與應用新冠肺炎疫情，加速全球產業數位轉...	0.008963	-0.002064	-0.007021	-0.025104	0.006794	-0.044399	0.078108	0.010230	-0.035018	0.002557	-0.048729	0.040761	0.039925	0.000723	0.027762	0.009268	-0.020337	0.008290	-(
166	2021-05-24	nan台股能否續彈或整理 仍須觀察疫情變化本土疫情擴散以來，台股走勢震盪，上週五反彈259....	0.022020	-0.019816	-0.005852	-0.003104	0.057515	-0.045100	0.080696	0.027528	-0.019105	0.037185	-0.023805	-0.005204	0.033009	0.010630	0.004874	0.028208	0.023649	-0.007868	(

LSTM

```
def LSTM_model(dates, prices, test_date, df):
    dataset = df.drop(columns=['date'])
    dataset = dataset.fillna(0)
    dataset = dataset.values

    # normalize the dataset
    scaler = MinMaxScaler(feature_range=(0, 1))
    scaler2 = MinMaxScaler(feature_range=(0, 1))

    data_scaler = scaler.fit_transform(numpy.array(dataset)[:,1:16])
    output_scaler = scaler2.fit_transform(numpy.array(dataset)[:,16].reshape(-1,1))

    dataset = numpy.concatenate((data_scaler, output_scaler), axis=1)

    # split into train and test sets
    train_size = len(dataset) - 2
    train, test = dataset[0:train_size, :], dataset[train_size:len(dataset), :]

    # reshape into X=t and Y=t+1
    look_back = 1
    trainX, trainY = create_dataset(train, look_back)
    testX, testY = create_dataset(test, look_back)
    X_train, X_test, y_train, y_test = train_test_split(trainX, trainY, test_size=0.33, random_state=42)
    # reshape input to be [samples, time steps, features]
    X_train = numpy.reshape(X_train, (X_train.shape[0], 1, X_train.shape[1]))
    X_test = numpy.reshape(X_test, (X_test.shape[0], 1, X_test.shape[1]))
    testX = numpy.reshape(testX, (testX.shape[0], 1, testX.shape[1]))
```

長短期記憶
(Long Short-Term Memory)

深度學習的一種。主要是以短期表現去預測未來長期會發生的事，適合用在股票預測、天氣預測等。

SVR

```
def SVR_rbf(dates, prices, test_date, df):
    svr_rbf = SVR(kernel='rbf', C=1e3, gamma=0.1)
    trainX, trainY, testX, testY = create_preprocessed_Dataset(df)
    # trainX = [item for sublist in trainX for item in sublist]
    # testX = [item for sublist in testX for item in sublist]
    X_train, X_test, y_train, y_test = train_test_split(trainX, trainY, test_size=0.33, random_state=42)
    svr_rbf.fit(trainX, trainY)
    decision_boundary = svr_rbf.predict(trainX)
    y_pred = svr_rbf.predict(X_test)
    test_score = mean_squared_error(y_test, y_pred)
    prediction = svr_rbf.predict(testX)[0]

    return (decision_boundary, prediction, test_score)
```

支持向量迴歸
(Support Vector Regression)

支持向量機(SVM)的延伸，不一樣的是SVR可以用來處理回歸的問題。SVR有分為線性與非線性，我們這次是以非線性的高斯轉換來做預測。

Linear Regression

```python
def linear_regression(dates, prices, test_date, df):
    lin_reg = LinearRegression()
    trainX, trainY, testX, testY = create_preprocessed_Dataset(df)
    # trainX = [item for sublist in trainX for item in sublist]
    # testX = [item for sublist in testX for item in sublist]
    X_train, X_test, y_train, y_test = train_test_split(trainX, trainY, test_size=0.33, random_state=42)
    lin_reg.fit(trainX, trainY)
    decision_boundary = lin_reg.predict(trainX)
    y_pred = lin_reg.predict(X_test)
    test_score = mean_squared_error(y_test, y_pred)
    prediction = lin_reg.predict(testX)[0]
    prediction = prediction[0]

    return (decision_boundary, prediction, test_score)
```

線性迴歸

依照餵給機器的訓練資料找出一個最接近各點線性方程式，並用函式去預測出新的資料(測試資料)應該在的位置。

Random Forest

```
def random_forests(dates, prices, test_date, df):
    rand_forst = RandomForestRegressor(n_estimators=10, random_state=0)
    trainX, trainY, testX, testY = create_preprocessed_Dataset(df)
    # trainX = [item for sublist in trainX for item in sublist]
    # testX = [item for sublist in testX for item in sublist]
    X_train, X_test, y_train, y_test = train_test_split(trainX, trainY, test_size=0.33, random_state=42)
    rand_forst.fit(trainX, trainY)
    decision_boundary = rand_forst.predict(trainX)
    y_pred = rand_forst.predict(X_test)
    test_score = mean_squared_error(y_test, y_pred)
    prediction = rand_forst.predict(testX)[0]

    return (decision_boundary, prediction, test_score)
```

隨機森林

屬於整體機器學習演算法(Ensemble learning)，一種由多個決策數組合而成，特徵是效能與預測能力很好，以結合多個弱學習器來建構一個強穩的模型，使模型較不會發生偏差。

KNN

```
def KNN(dates, prices, test_date, df):
    knn = KNeighborsRegressor(n_neighbors=3)
    trainX, trainY, testX, testY = create_preprocessed_Dataset(df)
    # trainX = [item for sublist in trainX for item in sublist]
    # testX = [item for sublist in testX for item in sublist]
    X_train, X_test, y_train, y_test = train_test_split(trainX, trainY, test_size=0.33, random_state=42)
    knn.fit(trainX, trainY)
    decision_boundary = knn.predict(trainX)
    y_pred = knn.predict(X_test)
    test_score = mean_squared_error(y_test, y_pred)
    prediction = knn.predict(testX)[0]
    prediction = prediction[0]

    return (decision_boundary, prediction, test_score)
```

K-近鄰演算法
(K-Nearest Neighbor)

輸入一筆資料 i 時，根據與 i 最鄰近的 k 個已知資料類別，去決定該筆輸入資料的類別。這種演算法除了做分類用途，也能用來做近似項目的搜尋。

03
PART

未來展望

未來展望

自動化爬取資料

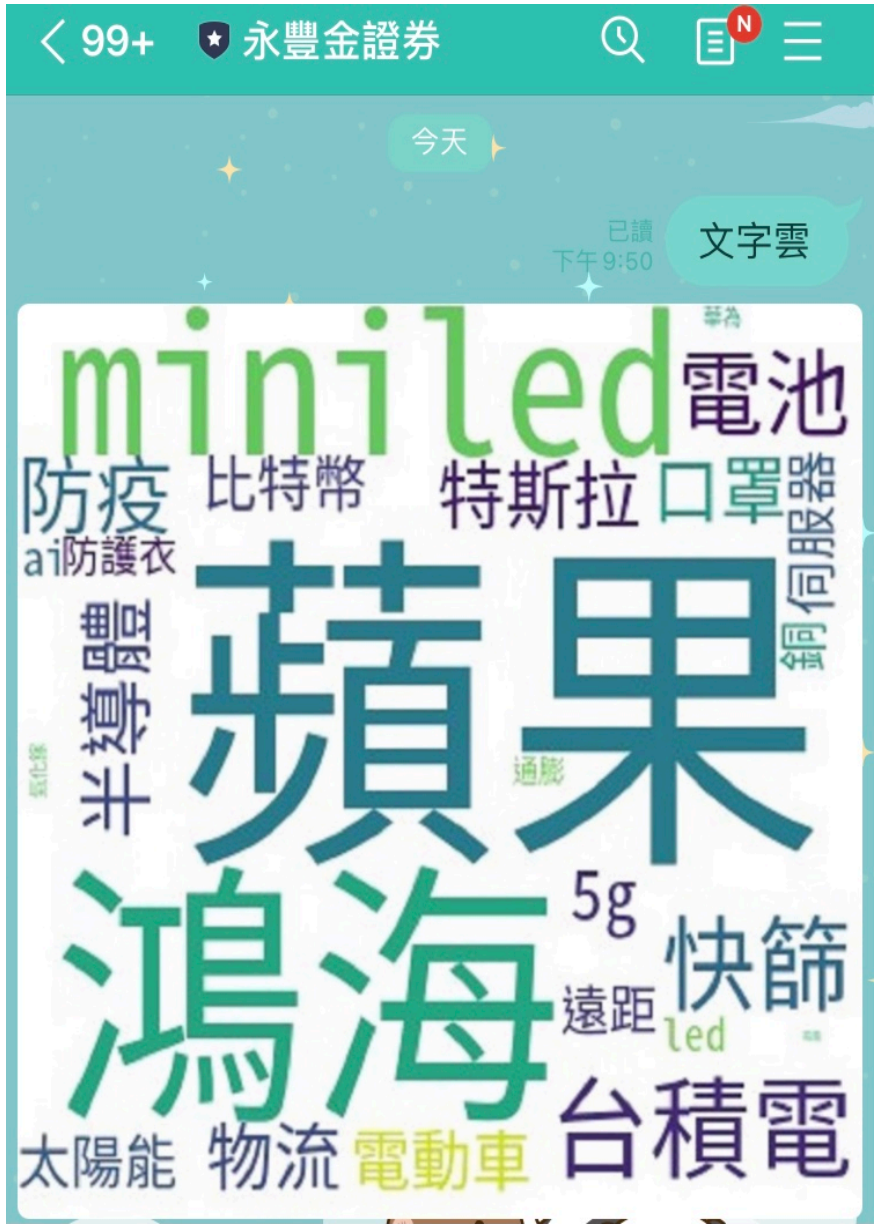
評分機制

增加模型準確率

04
PART

成果展示

成果展示



代號	股名	收盤價	漲幅(%)	成交量
2340	光磊	25.25	-8.51	23860
2393	億光	48.8	-1.01	13039
2409	友達	23.05	1.32	230660
2426	鼎元	26.05	0	1695
3038	全台	20.8	-0.24	966
3049	和鑫	12.75	0	3438
3481	群創	21.35	0.95	329481
4934	太極	23.9	0.42	803
4956	光鋐	22.75	-0.44	664
6116	彩晶	22.75	4.12	172261

更新日期：2021/06/11　資料僅供參考，以往績效不保證未來獲利

成果展示

股票名稱 / 代碼	毛利連續成長季數	營收連續成長季數	營業利益連續成長季數	稅後淨利連續成長季數	最近一季毛利	最近一季營業利益	最近一季稅後淨利	最近一季營收
友達(2409)	50%	60%	60%	60%	100%	64%	73%	73%
彩晶(6116)	75%	60%	60%	60%	91%	91%	100%	100%
鼎元(2426)	25%	20%	60%	80%	64%	82%	36%	64%
宏齊(6168)	100%	80%	80%	80%	45%	45%	55%	18%
台表科(6278)	75%	40%	20%	60%	55%	18%	45%	55%
光磊(2340)	75%	40%	20%	20%	27%	36%	27%	45%
太極(4934)	0%	0%	0%	0%	0%	0%	0%	0%
和鑫(3049)	25%	20%	20%	20%	9%	9%	18%	9%
億光(2393)	75%	80%	80%	80%	36%	100%	91%	36%
全台(3038)	25%	20%	80%	20%	18%	27%	9%	27%
光鋐(4956)	75%	100%	100%	100%	73%	55%	64%	91%

成果展示

永豐金證券 SinoPac Securities

股票分析

股票名稱 / 代碼

友達(2409)

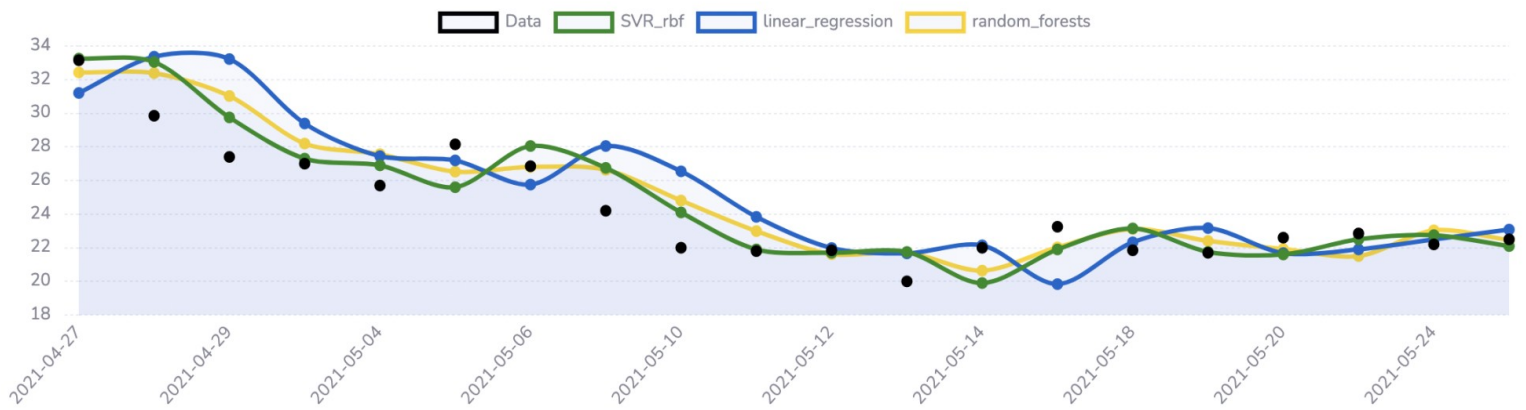
預測模型

Nothing selected

Submit

友達(2409)

股票預測趨勢



股票預測結果

可參考上述股票預測趨勢圖，選擇最適合您的股價波動，並取得預測之收盤價。

預測日期： 2021-05-28

Predictions

Model	收盤價

成果展示

股票預測結果

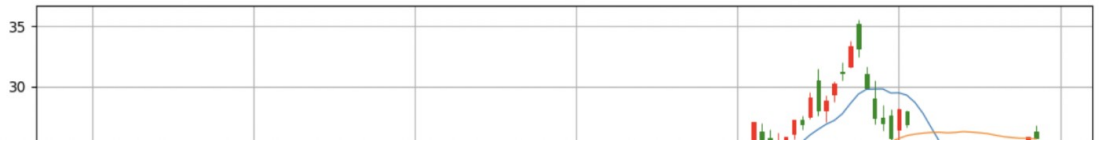
可參考上述股票預測趨勢圖，選擇最適合您的股價波動，並取得預測之收盤價。
預測日期： 2021-05-28

Predictions

Model	收盤價
Original	23.5
SVR_rbf	11.983175204687324
linear_regression	22.596895770231086
random_forests	22.340000343322753

財報分析

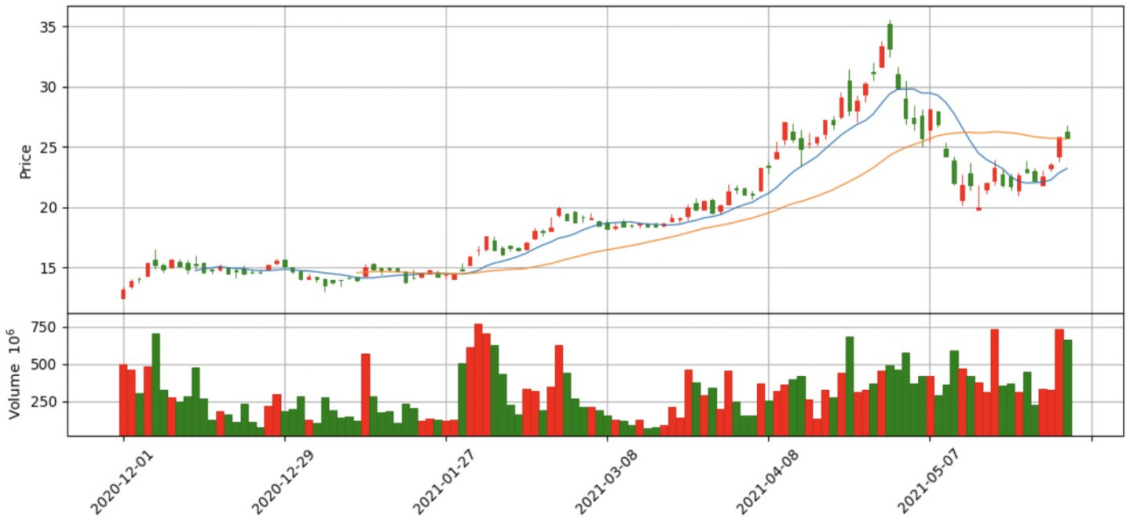
K線圖



成果展示

財報分析

K線圖



流動資產與負債佔比

2020

流動資產(單位:千元)

2020

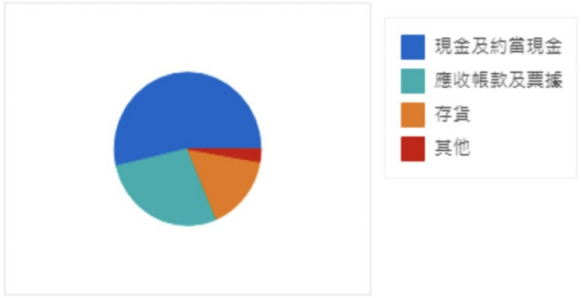
負債(單位:千元)

成果展示

流動資產與負債佔比

2020

流動資產(單位:千元)



流動資產可以判斷公司的償還能力，如果流動比率太低需注意營運資金不足的問題。

2020

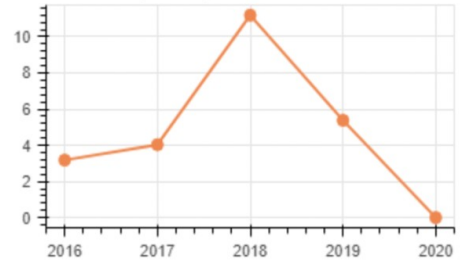
負債(單位:千元)

負債可以判斷公司的償債能力，如果流動比率太低需注意營運資金不足的問題。

PE Ratio & 現金殖利率

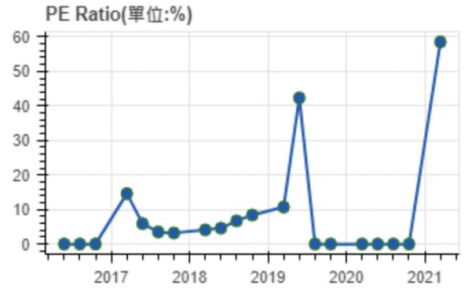
PE Ratio(單位:%)

現金殖利率(單位:%)



成果展示

PE Ratio & 現金殖利率



PE Ratio（本益比）：本益比可判斷目前公司股價的價位合理性，可能代表目前股價被相對高估，但也可能代表該股為目前熱門趨勢股、或未來成長性受投資人看好。

現金殖利率：殖利率為將錢投入個股後所能獲得的現金股利比率，追求穩定配息的投資人可利用此指標搭配營收、毛利率的穩定度判斷是否為獲利穩健、可持續創造現金流之公司。

營收 & 單季營收年增長率

工作分配

組員	專案製作
戴緯綾	擔任組長、業師聯繫、網路爬蟲、PPT製作、影片錄製
陳首暄	主題發想、網路爬蟲、機器學習、網頁製作
李重諺	網路爬蟲、情緒分析、文字雲製作
何佳穎	網頁刻板、網頁製作、Notion管理、資料視覺化
陳若昕	主題發想、網頁規劃、財務數據蒐集、Notion管理、影片錄製

永豐金證券
SinoPac Securities

謝謝觀看