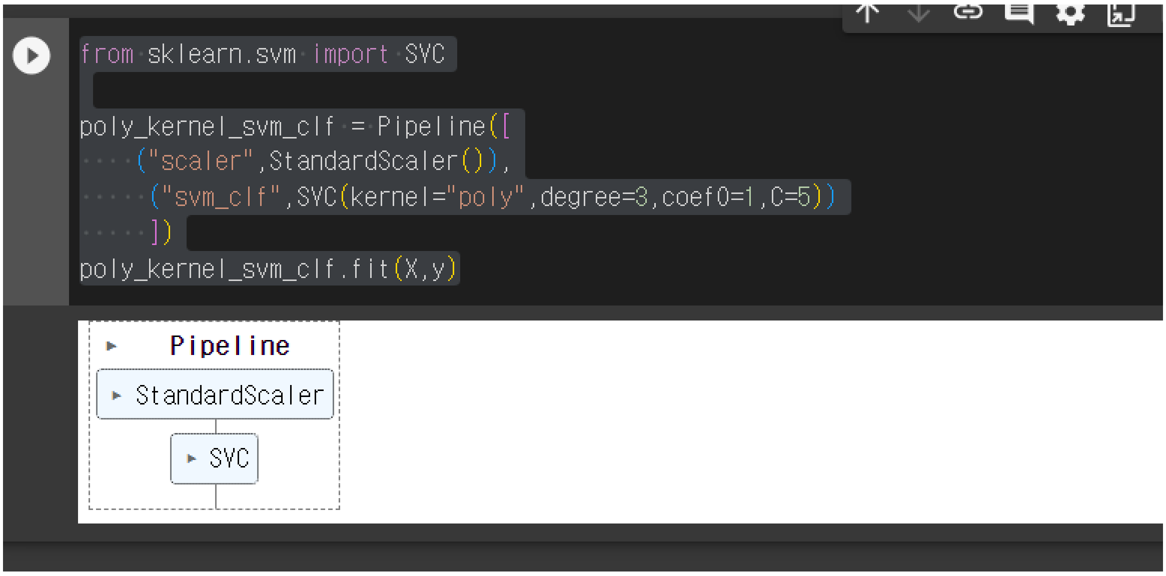
기계학습_231010

▼ 실행시 결과값이 이상하게 나와서 다시 실습 해봐야 할듯함



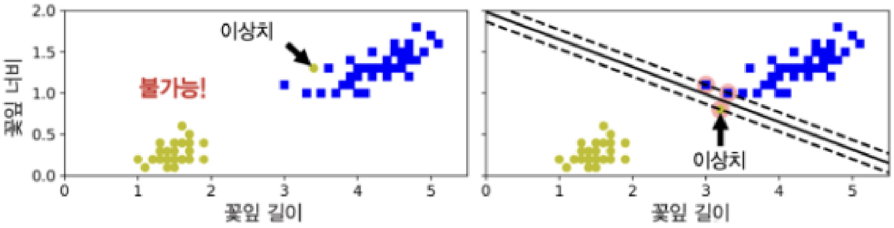
▼ 찾아둔 곳(동일한 코드는 아닌듯 중간중간 다름)

https://colab.research.google.com/github/codingalzi/handson-ml2/blob/master/notebooks/handson-ml2-05.ipynb#scrollTo=LcXe3Soa_6bR

SVM

5.1.1 하드/소프트 마진 분류

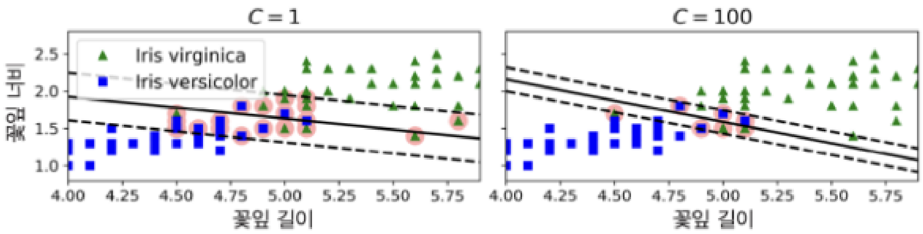
-하드 마진 분류(도로로 완벽 구분) : 이상치에 민감, 데이터가 선형적으로 구분 되어야함



▲ 이상치에 민감한 하드 마진

	왼편 그래프	오른편 그래프
이상치:	타 클래스에 섞임	타 클래스에 매우 가까움
하드 마진 분류:	불가능	가능하지만 일반화 어려움

-소프트 마진 분류(도로폭 넓게 유지 및 마진오류 사이의 균형 잡기) :



5.1.2 선형 SVM 지원 모델 예제

(코딩)

→ 힌지 개념이 중요

5.2 비선형 SVM 분류

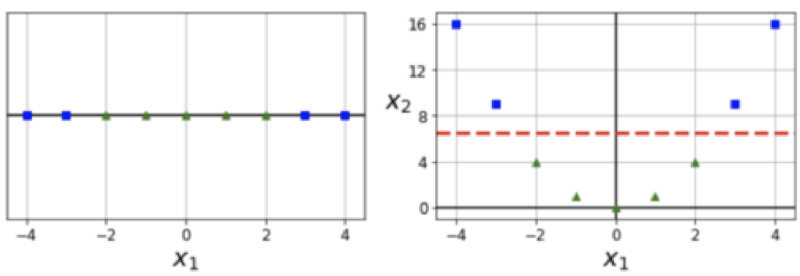
-선형 SVM에 특성 추가 : 다항 특성 활용, 유사도 특성활용

-SVC+커널 트릭(특성 추가 않하면서 특성 많이추가한것같은 결과얻음) : 다항식 커널, 가우시안 RBF 커널

5.2.1 선형 SVM+ 다항 특성 추가

-특성 a하나만 갖는 모델에 새로운 특성 b추가한 후 분류

ex) 기존 특성 : x1, 추가 특성 (x1)^2



5.2.2 유사도 특성 활용

-잘 안보이는 데이터(좌)를 유사도 함수를 이용해서 바꿔서(우) 분류하려고함

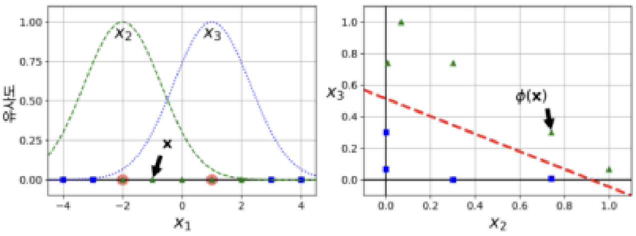
→ 차원이 커지면서 선형적으로 구분될 가능성이 높아짐

-유사도함수 : 가우시안 방사 기저 함수(RBF)

$$\phi(\mathbf{x}, \ell) = \exp(-\gamma \|\mathbf{x} - \ell\|^2)$$

ℓ: 랜드마크

γ: 랜드마크에서 멀어질 수록 0에 수렴하는 속도를 조절함



→랜드마크 : -2, 1

→ex)위의 랜드마크 예시(-2,1)로 계산시 0.74, 0.30가 나옴(우측 사진)

5.2.3 다항식 커널

-머신러닝에서의 커널 : 특정 함수를 곱한다,내적한다

-간단하지만 낮은차수일시 매우 복잡한 데이터셋을 잘 표현하지 못함 높은 차수일시 모델이 느려짐

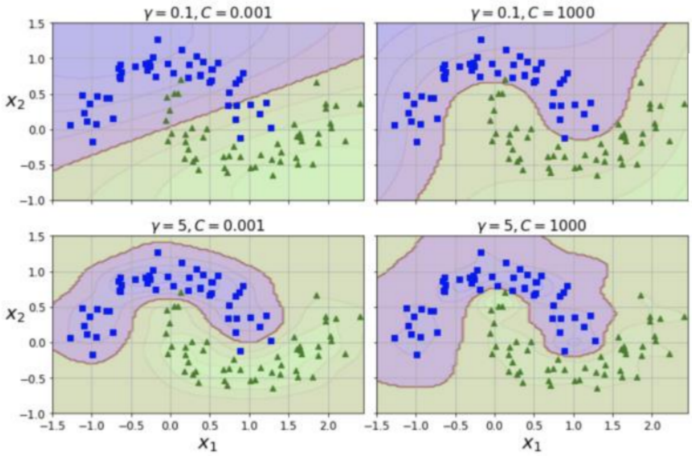
<중간고사에 다음을 외우는거 나올것임>

```
from sklearn.svm import SVC

poly_kernel_svm_clf = Pipeline([("scaler",StandardScaler()),("svm_clf",SVC(kernel="poly",degree=3,coef0=1,C=5))])
poly_kernel_svm_clf.fit(X,y)
```

5.2.4 가우시안 RBF 커널

-하이퍼 파라미터 ‘감마’ : 증가시 결정경계 구불구불해지고 감소시 경계가 부드러워짐



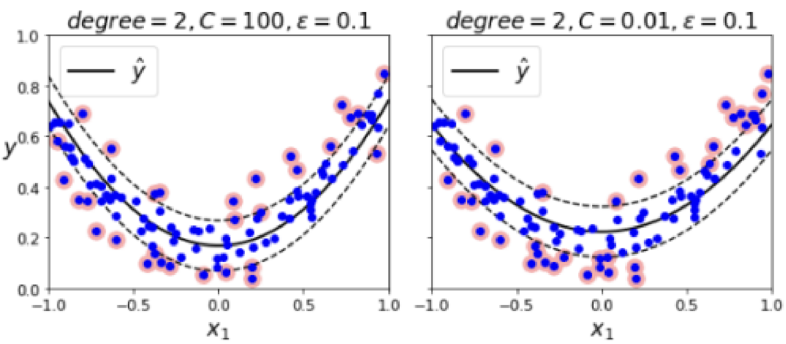
	상단 왼편 그래프	상단 오른편 그래프
kernel="rbf":	가우시안 RBF 커널	가우시안 RBF 커널
gamma=ga:	ga=0.1: 랜드마크에 조금 집중	ga=0.1: 랜드마크에 조금 집중
C=C:	C=0.001: 가중치 규제 많이	C=1000: 가중치 규제 적게
	따라서 마진 도로폭 넓게	따라서 마진 도로폭 좁게

	하단 왼편 그래프	하단 오른편 그래프
kernel="rbf":	가우시안 RBF 커널	가우시안 RBF 커널
gamma=ga:	ga=5.0: 랜드마크에 많이 집중	ga=5.0: 랜드마크에 많이 집중
C=C:	C=0.001: 가중치 규제 많이	C=1000: 가중치 규제 적게
	결정 경계 덜 민감	결정 경계가 보다 민감

▲ RBF 커널을 사용한 SVM 분류기

5.3 svm 회귀

차이점이 확연히 보이지는 않음..ㅠㅠ



▲ 그림 5-11 2차 다항 커널을 사용한 SVM 회귀

5.4 SVM 이론

-선형 SVM 작동원리

-커널 SVM 작동원리

-온라인 SVM

5.4.1 결정 함수와 예

<못적어땅>

5.4.2 목적함수

-하드마진SVM 분류기의 목적함수 :

$$\frac{1}{2}\mathbf{w}^T\mathbf{w}$$

→ 목적함수를 최소화 시키는 w,b 구하기

$$t^{(i)}(\mathbf{w}^T\mathbf{x}^{(i)} + b) \geq 1$$

$x^{(i)}$: i 번째 샘플
$t^{(i)}$: 양성 샘플일 때 1, 음성 샘플일 때 -1

-소프트마진SVM 분류기의 목적함수 :

$$\frac{1}{2}\mathbf{w}^T\mathbf{w} + C\sum_{i=0}^{m-1}\zeta^{(i)}$$

→ 마진오류가 하나 더 들어감

→ 목적함수를 최소화 시키는 w,b 구하기

$$t^{(i)}(\mathbf{w}^T\mathbf{x}^{(i)} + b) \geq 1 - \zeta^{(i)}$$

: $x^{(i)}$: i 번째 샘플
$t^{(i)}$: 양성 샘플일 때 1, 음성 샘플일 때 -1
$\zeta^{(i)} \geq 0$: **슬랙 변수**. i 번째 샘플이 얼마나 마진을 위반할지 정함.

5.4.3 <책 범위 벗어난 부분>

5.4.4 커널 SVM

-2차 다항 커널 작동 아이디어

→원래식

$$\frac{1}{2}\sum_{i=0}^{m-1}\sum_{j=0}^{m-1}\alpha^{(i)}\alpha^{(j)}t^{(i)}t^{(j)}\boxed{\phi(\mathbf{x}^{(i)})^T\phi(\mathbf{x}^{(j)})} - \sum_{i=0}^{m-1}\alpha^{(i)}$$

→변환된벡터점곱= 원래벡터점곱

$$\begin{aligned}\phi(\mathbf{a})^T\phi(\mathbf{b}) &= \begin{pmatrix} {a_1}^2 \\ \sqrt{2}\,a_1a_2 \\ {a_2}^2 \end{pmatrix}^T \begin{pmatrix} {b_1}^2 \\ \sqrt{2}\,b_1b_2 \\ {b_2}^2 \end{pmatrix} = {a_1}^2{b_1}^2 + 2a_1b_1a_2b_2 + {a_2}^2{b_2}^2 \\ &= (a_1b_1 + a_2b_2)^2 = \left(\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}^T \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}\right)^2 = (\mathbf{a}^T\mathbf{b})^2\end{aligned}$$

→ 결과적으로 이 함수에 대한 최적화 문제를 해결

$$\frac{1}{2}\sum_{i=0}^{m-1}\sum_{j=0}^{m-1}\alpha^{(i)}\alpha^{(j)}t^{(i)}t^{(j)}\boxed{\left(\mathbf{x}^{(i)^T}\mathbf{x}^{(j)}\right)^2}-\sum_{i=0}^{m-1}\alpha^{(i)}$$

-머신러닝에서의 커널 : φ모르더라도 원래 백터 a와 b에 기반한 점곱을 계산할 수 있는 함수

5.4.5 커널 SVM

-선형 온라인 SVM

식 5-13: 선형 SVM 분류기의 비용 함수

$$J(\mathbf{w},b)=\frac{1}{2}\mathbf{w}^T\mathbf{w}+C\sum_{i=1}^{m}max\left(0,t^{(i)}-(\mathbf{w}^T\mathbf{x}^{(i)}+b)\right)$$