Метод опорных векторов (support vector machines)

PIERIAN DATA

Метод опорных векторов

- Метод опорных векторов (support vector machines) - это более сложный алгоритм, чем KNN, но всё начинается с простого вопроса:
  - Можем ли мы провести гиперплоскость, которая хорошо отделит классы друг от друга?
- Чтобы ответить на этот вопрос, давайте сначала взглянем на историю развития метода опорных векторов.

Метод опорных векторов

- В этом разделе курса мы изучим:
  - Историю метода опорных векторов
  - Теорию метода
  - Пример для задач классификации
  - Пример для задач регрессии
  - Проект для проверки навыков и разбор решений

Метод опорных векторов (support vector machines)

История

Метод опорных векторов

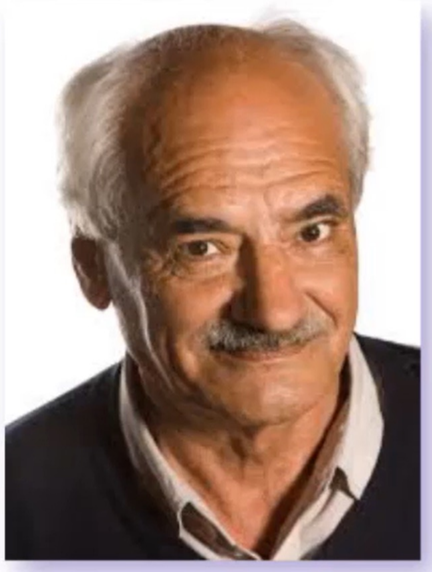
- Метод опорных векторов (support vector machines) – один из относительно недавно изобретённых алгоритмов машинного обучения, который мы изучим.
- Давайте кратко взглянем на его историю...

Метод опорных векторов

- 1960-е годы: Владимир Вапник защитил кандидатскую диссертацию в Институте проблем управления (Москва)
- Работал в этом институте 1961-1990

Метод опорных векторов

- 1963: Владимир Вапник и Алексей Червоненкис опубликовали “метод обобщённого портрета” для анализа изображений компьютерами



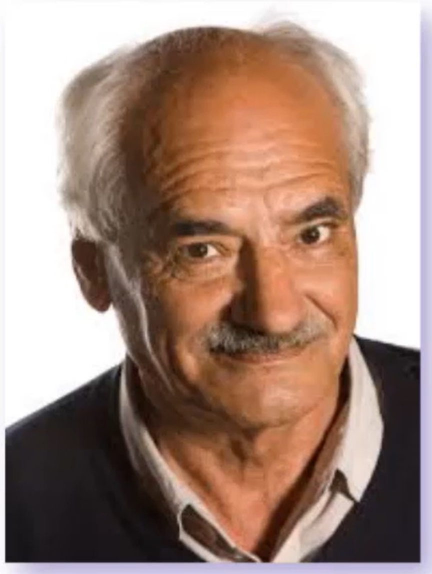
PIERIAN DATA

Метод опорных векторов

- 1964: М.А.Айзерман, М.М.Браверман, Л.И.Розоноэр публикуют “Теоретические основы метода потенциальных функций в задаче об обучении автоматов разделению входных ситуаций на классы”
- Даётся геометрическая интерпретация ядер (kernels) как произведений в пространстве признаков.

Метод опорных векторов

- 1974: Вапник и Червоненкис продолжают работу, публикуют книгу “Теория распознавания образов”



PIERIAN DATA

Метод опорных векторов

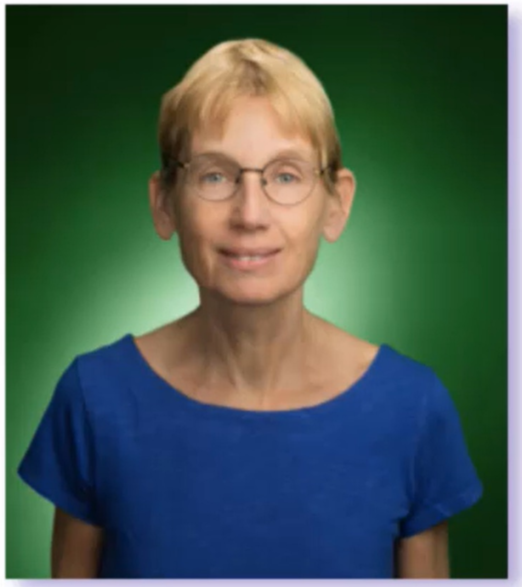
- 1960e-1990e года: Вапник продолжает работу над дальнейшим развитием метода опорных векторов

Метод опорных векторов

- 1992: Бернхард Бозер, Изабель Гюйон и Владимир Вапник предлагают нелинейный классификатор, применяя так называемый kernel trick для гиперплоскостей с максимальным зазором.

Метод опорных векторов

- 1995: Коринна Кортес и Владимир Вапник публикуют вариант метода опорных векторов с мягким зазором (soft margin)



Метод опорных векторов

- 1997: Владимир Вапник, Харрис Друкер, Кристофер Бургес, Линда Кауфман и Александр Смола публикуют работу “Support Vector Regression Machines”, расширяя диапазон применения SVM за пределы задач классификации.

Метод опорных векторов

- Чтобы получить ту версию метода опорных векторов, которую мы будем использовать, потребовалось более 30 лет!
- Сейчас продолжаются работы по увеличению производительности, а также теоретические исследования.

Метод опорных векторов

- Далее мы посмотрим теорию метода, постепенно продвигаясь вперед.
- Начнём с классификаторов с максимальным зазором (maximum margin classifiers), затем посмотрим классификатор опорных векторов и, наконец, сам метод опорных векторов.

Метод опорных векторов (support vector machines)

Теория – гиперплоскости и зазоры

Метод опорных векторов

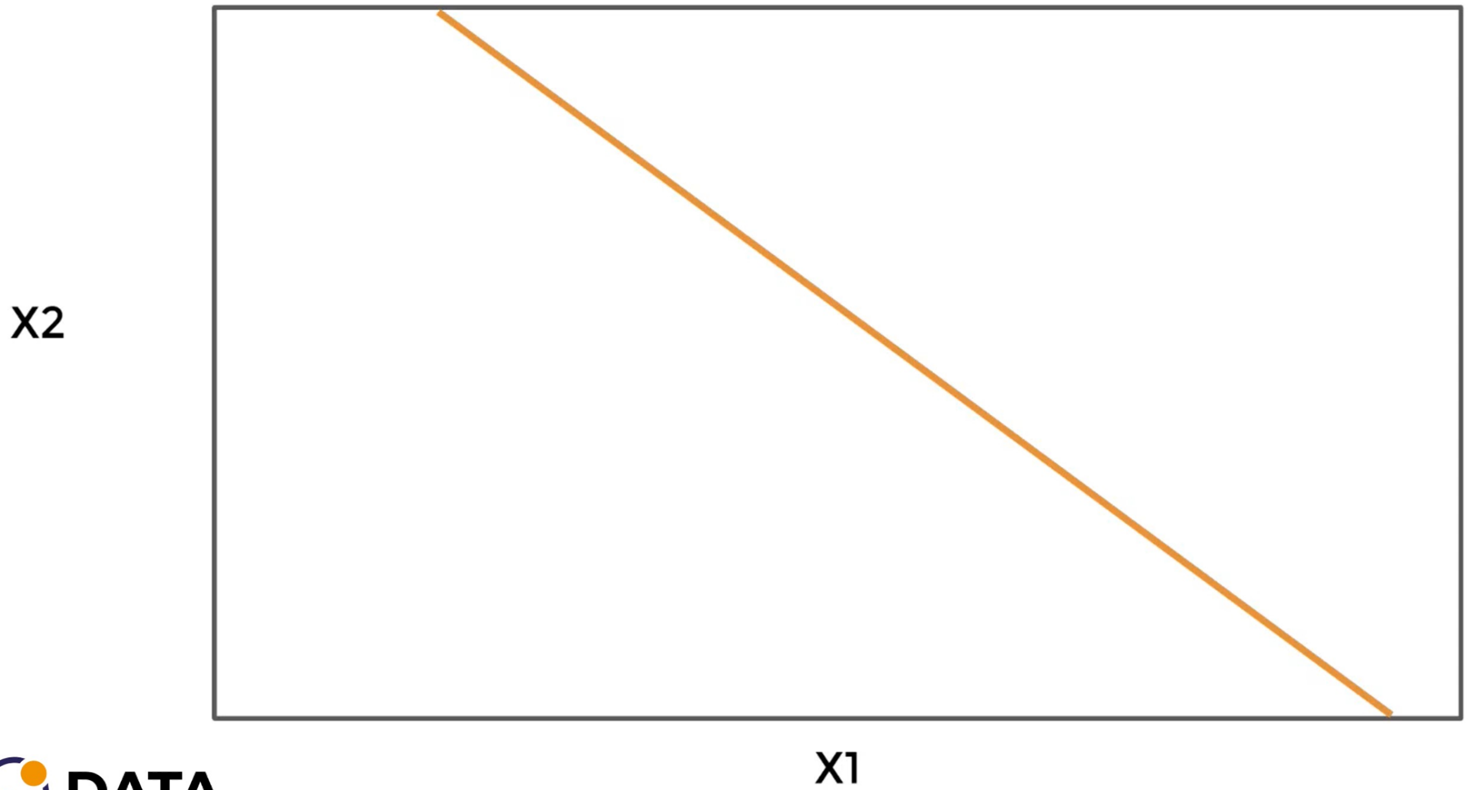
- Начнём с термина гиперплоскость. В N-мерном пространстве гиперплоскость – это подпространство размерности N-1.
- Например:
  - В 1-мерном пространстве гиперплоскость – это точка
  - В 2-мерном пространстве гиперплоскость – это линия
  - В 3-мерном пространстве гиперплоскость – это плоская поверхность

Метод опорных векторов

- В 1-мерном пространстве - точка:

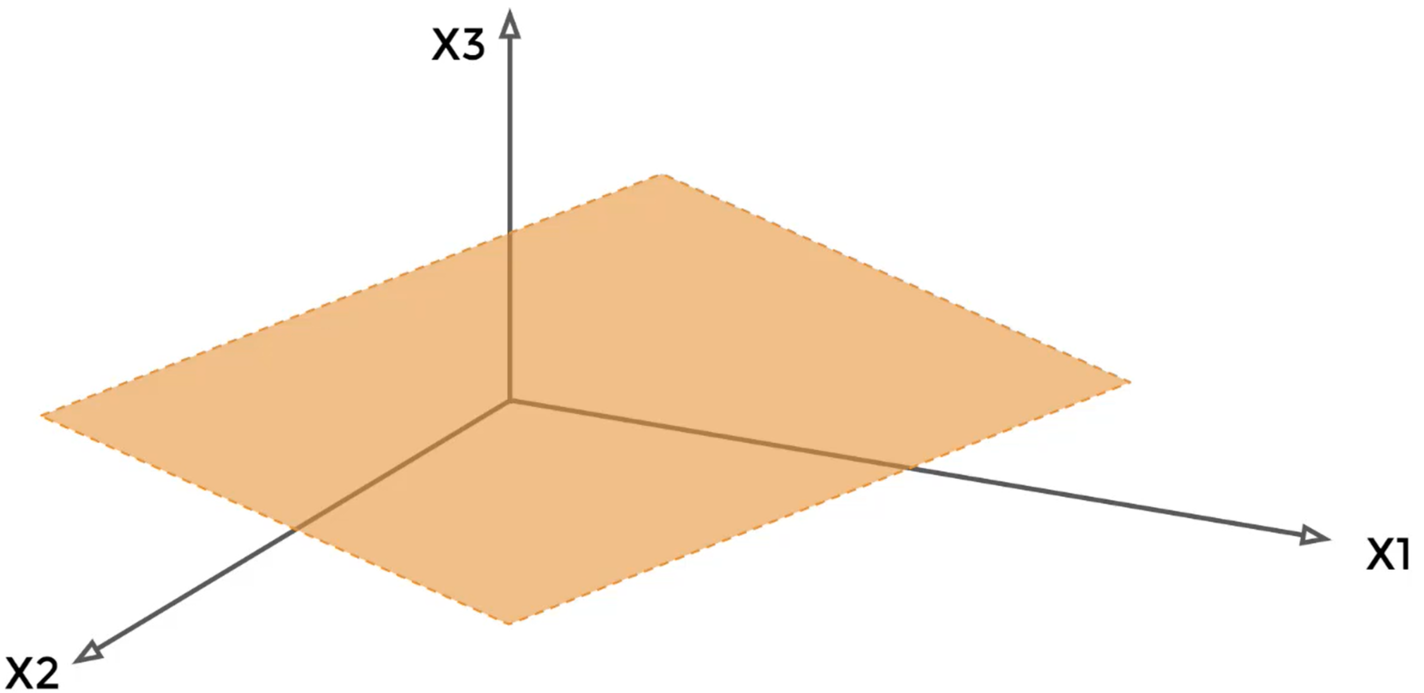
Метод опорных векторов

- В 2-мерном пространстве - линия:



Метод опорных векторов

- В 3-мерном пространстве – плоская поверхность:



Метод опорных векторов

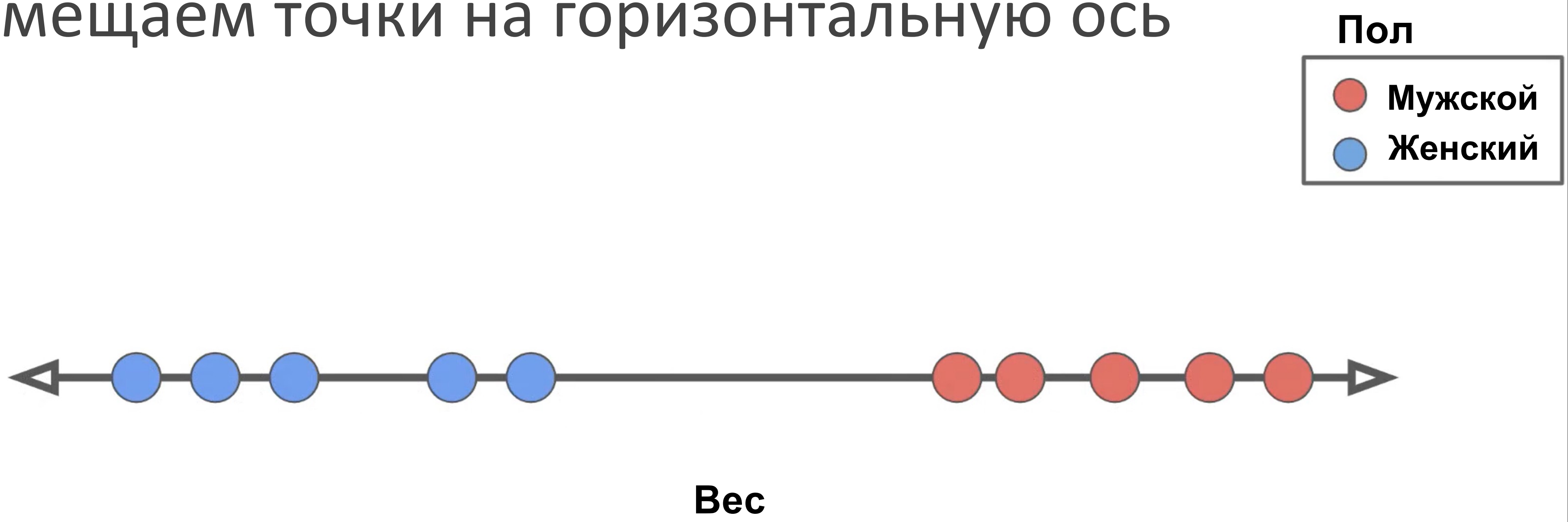
- Основная идея метода опорных векторов – в том, что мы можем найти гиперплоскость, отделяющую классы друг от друга.
- Когда новые точки оказываются по ту или иную сторону этой гиперплоскости, мы можем назначать этим точкам классы.

Метод опорных векторов

- Представим себе набор данных с одним признаком и бинарной целевой переменной.
- Например:
  - Признак - вес цыплят.
  - Целевая переменная – пол цыплят.
- Как это будет выглядеть визуально?

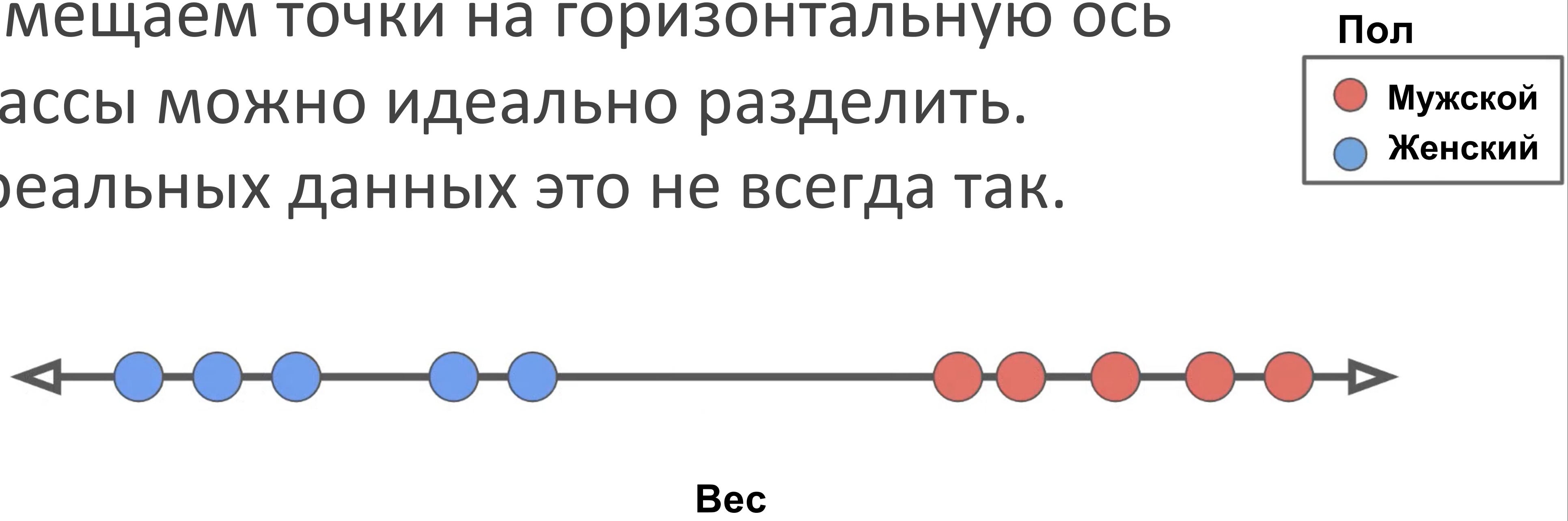
Метод опорных векторов

- Помещаем точки на горизонтальную ось



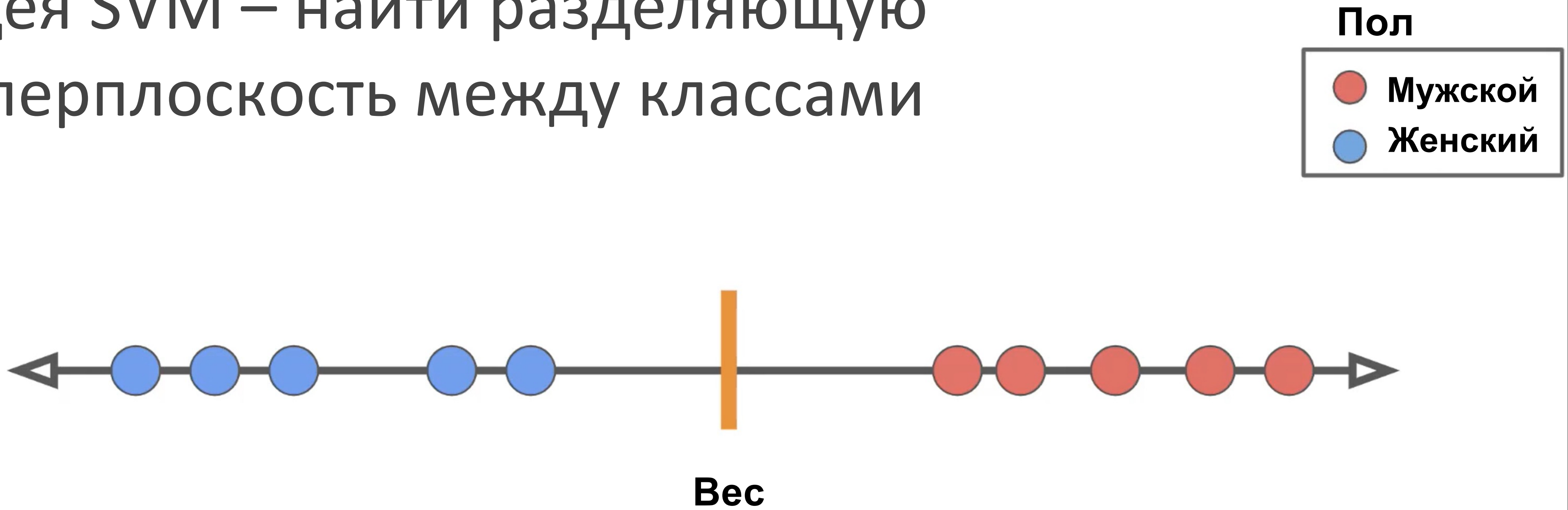
Метод опорных векторов

- Помещаем точки на горизонтальную ось
- Классы можно идеально разделить.
- В реальных данных это не всегда так.



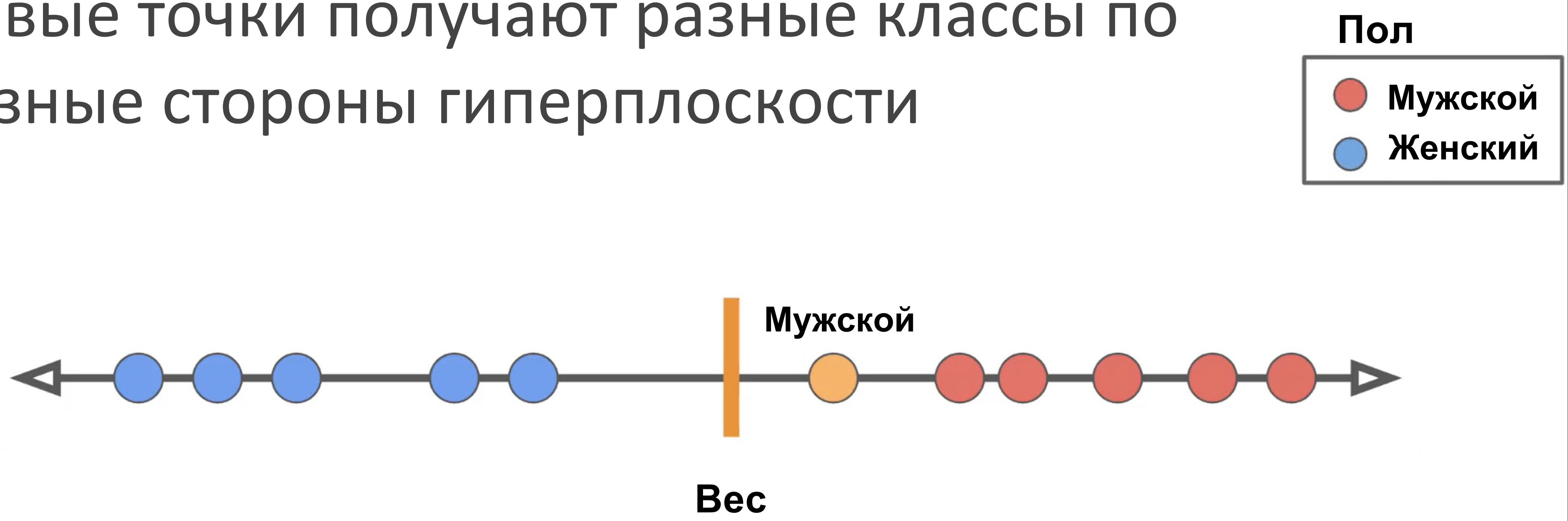
Метод опорных векторов

- Идея SVM – найти разделяющую гиперплоскость между классами



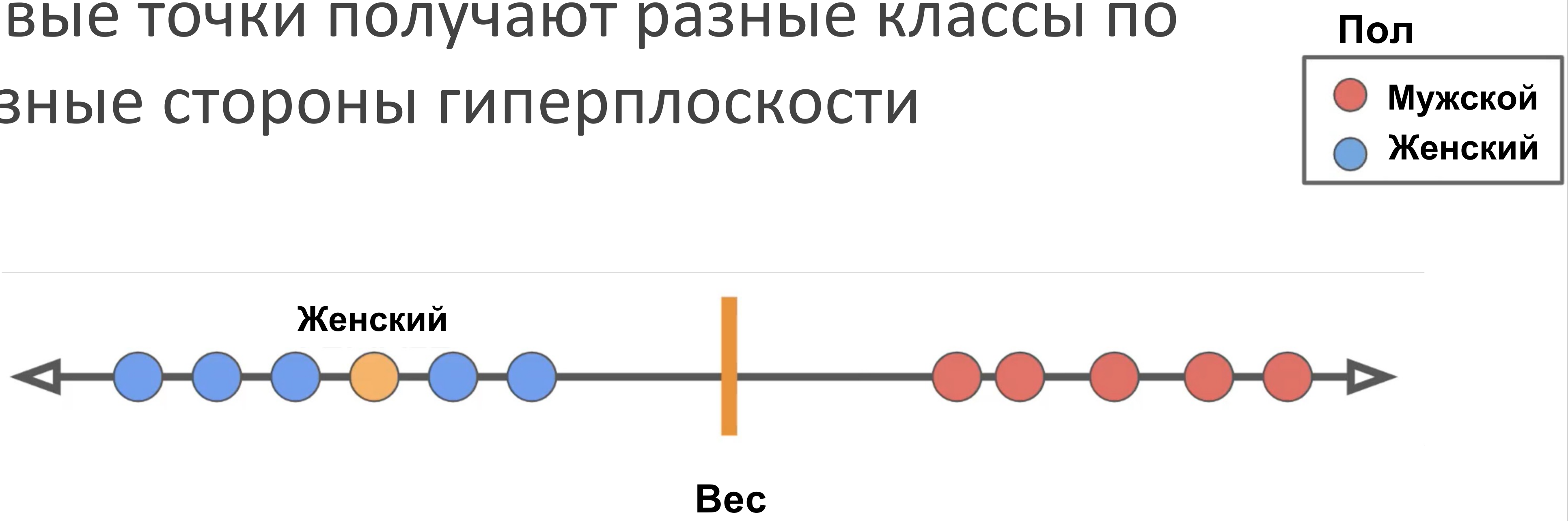
Метод опорных векторов

- Новые точки получают разные классы по разные стороны гиперплоскости



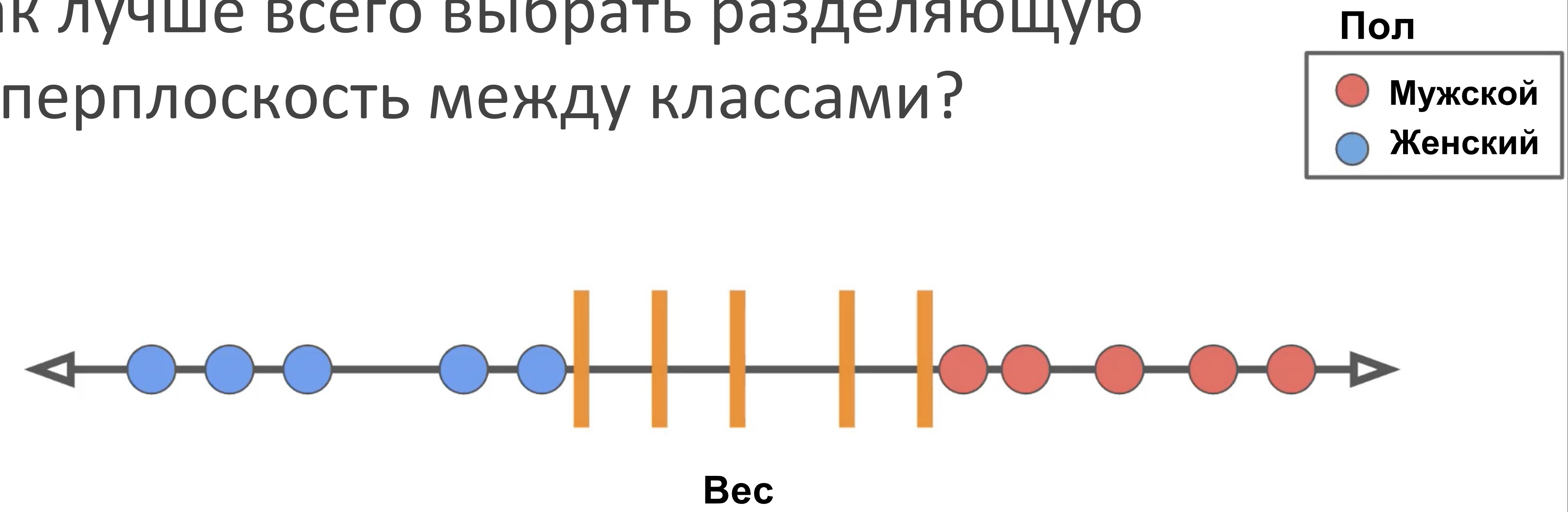
Метод опорных векторов

- Новые точки получают разные классы по разные стороны гиперплоскости



Метод опорных векторов

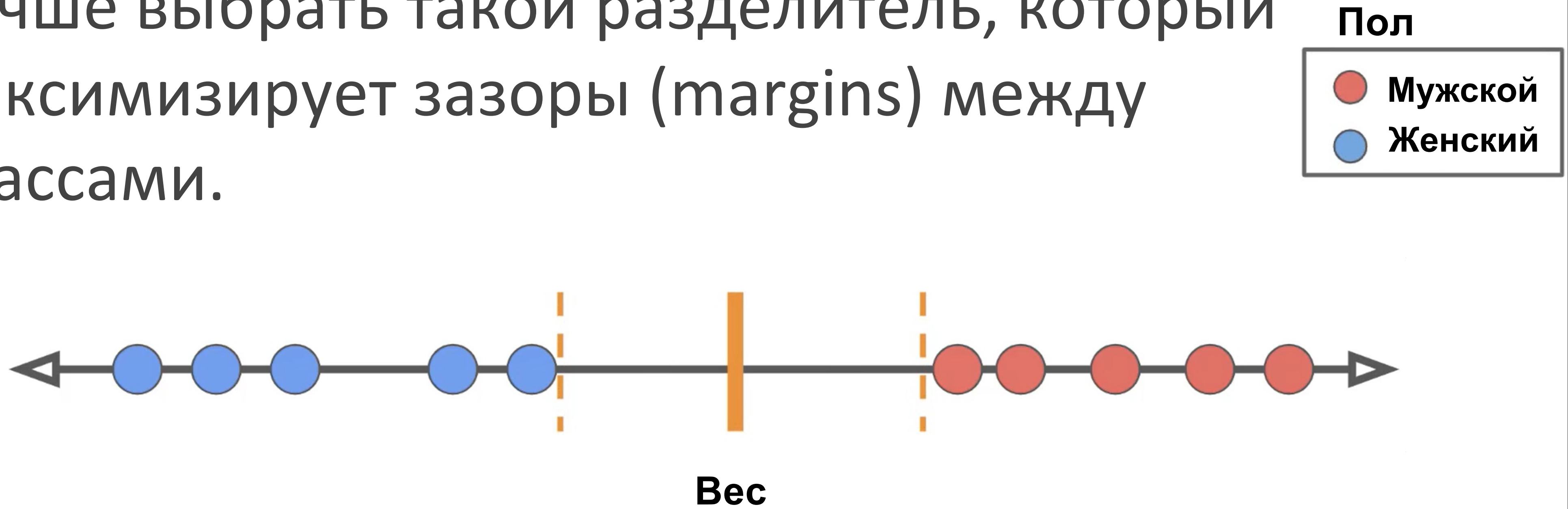
- Как лучше всего выбрать разделяющую гиперплоскость между классами?



PIERIAN DATA

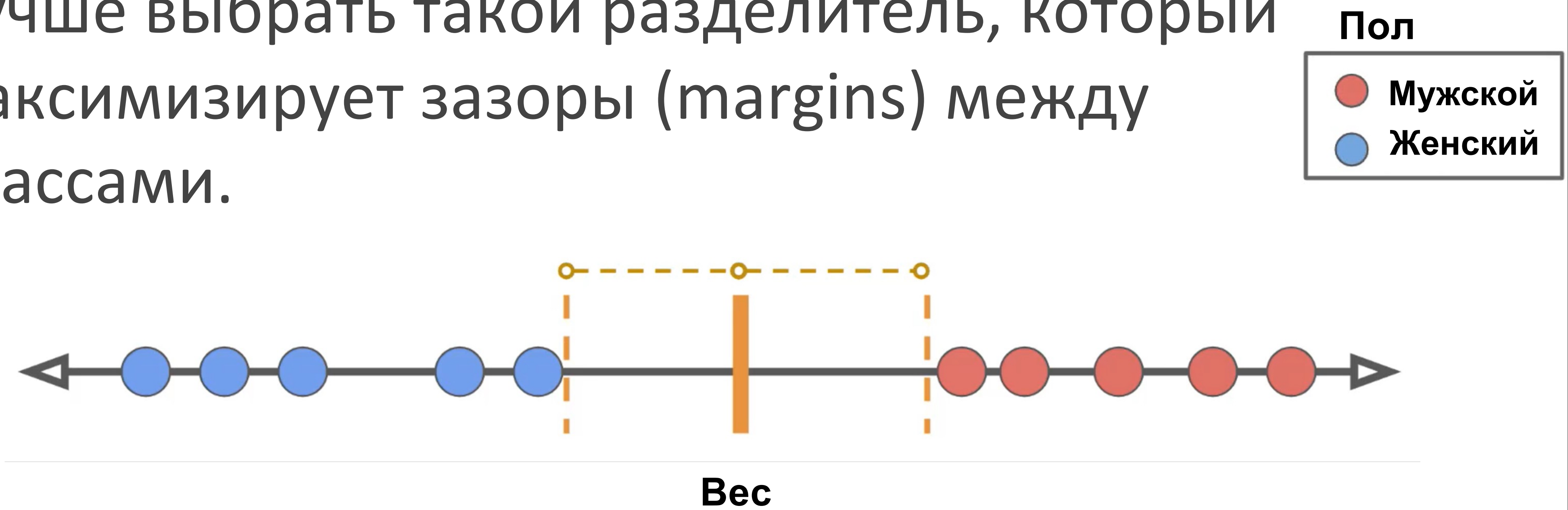
Метод опорных векторов

- Лучше выбрать такой разделитель, который максимизирует зазоры (margins) между классами.



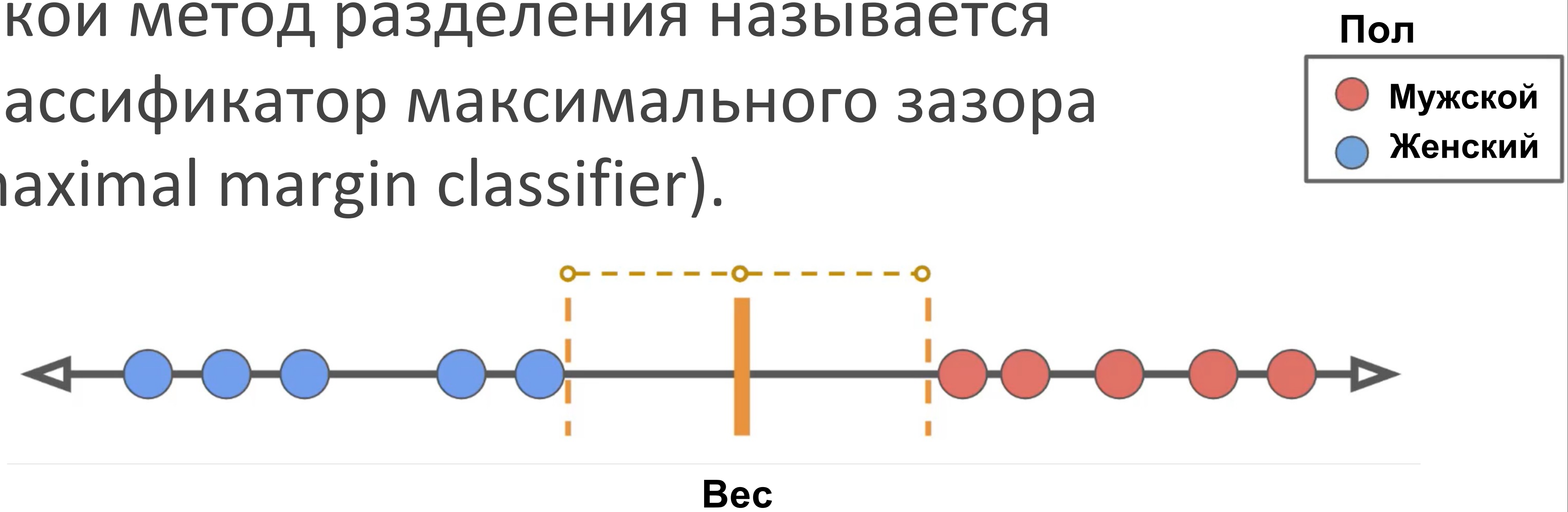
Метод опорных векторов

- Лучше выбрать такой разделитель, который максимизирует зазоры (margins) между классами.



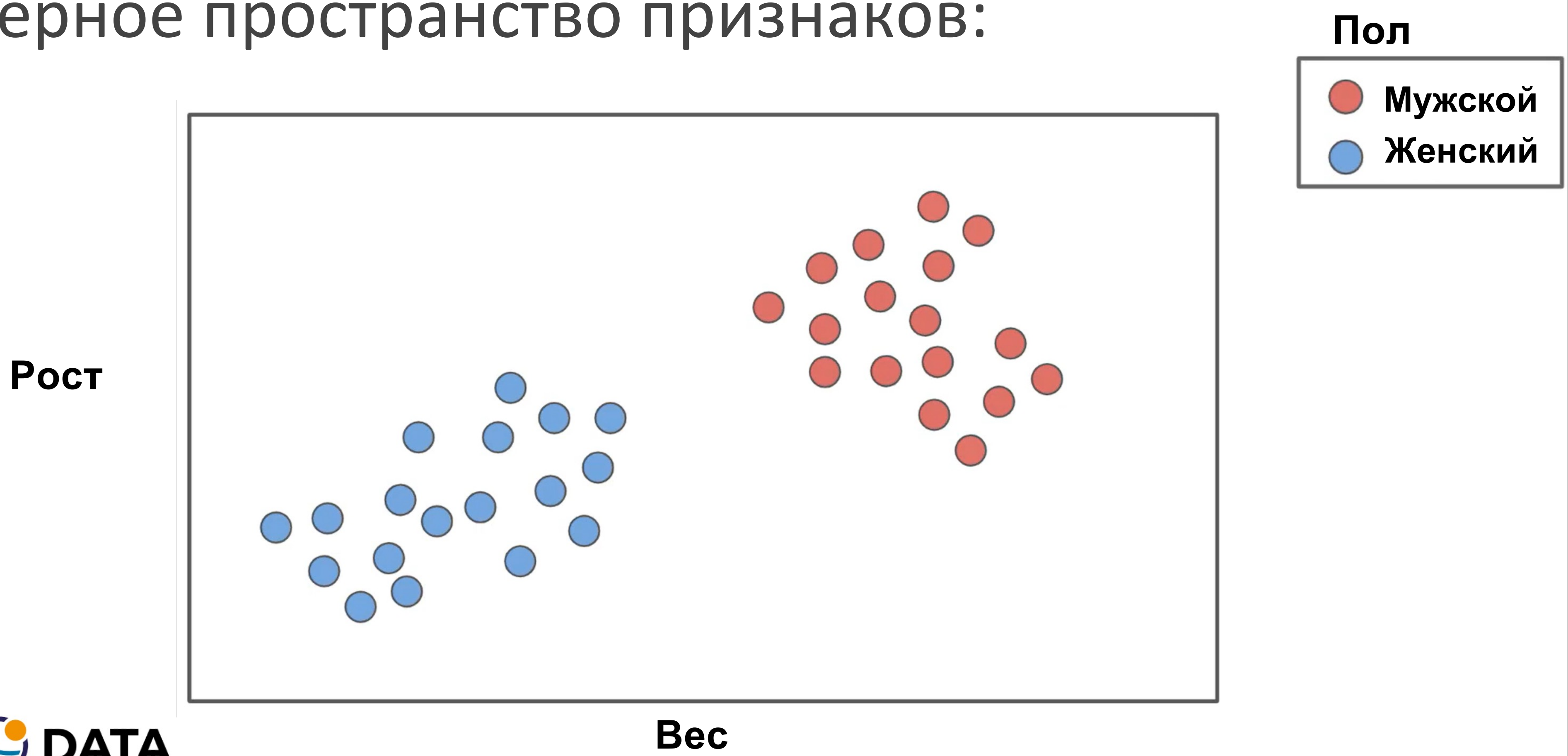
Метод опорных векторов

- Такой метод разделения называется классификатор максимального зазора (maximal margin classifier).



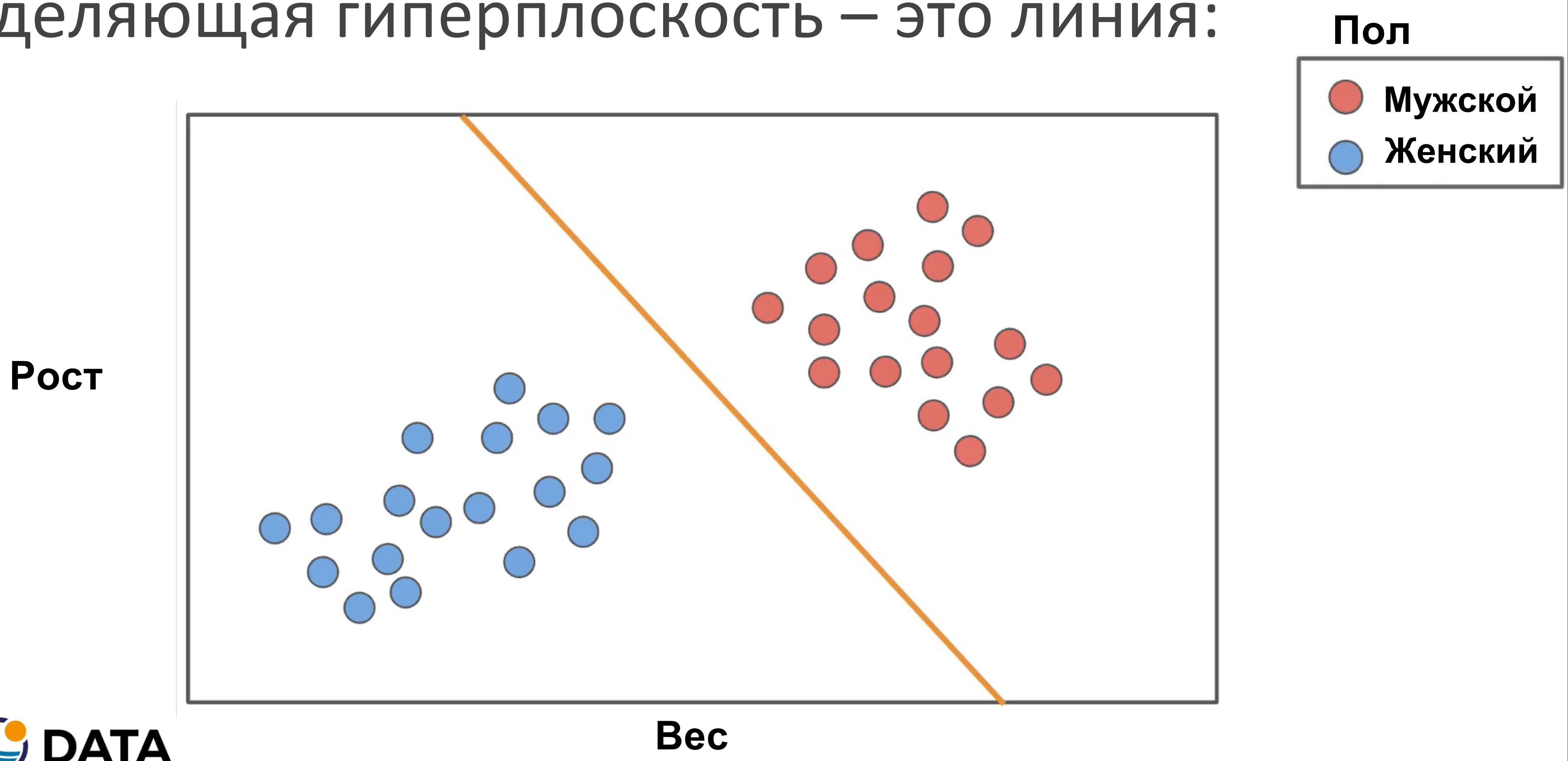
Метод опорных векторов

- 2-мерное пространство признаков:



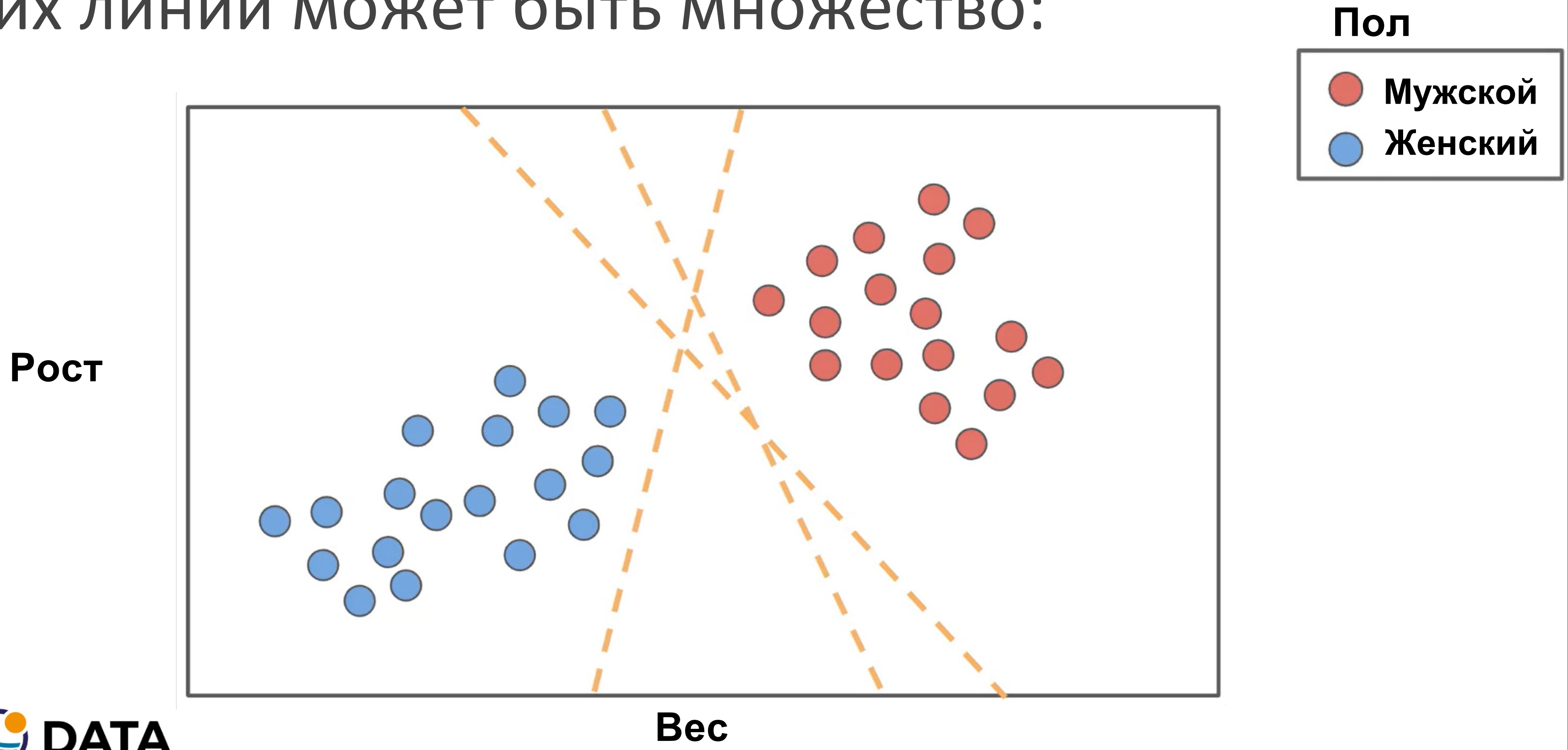
Метод опорных векторов

- Разделяющая гиперплоскость – это линия:



Метод опорных векторов

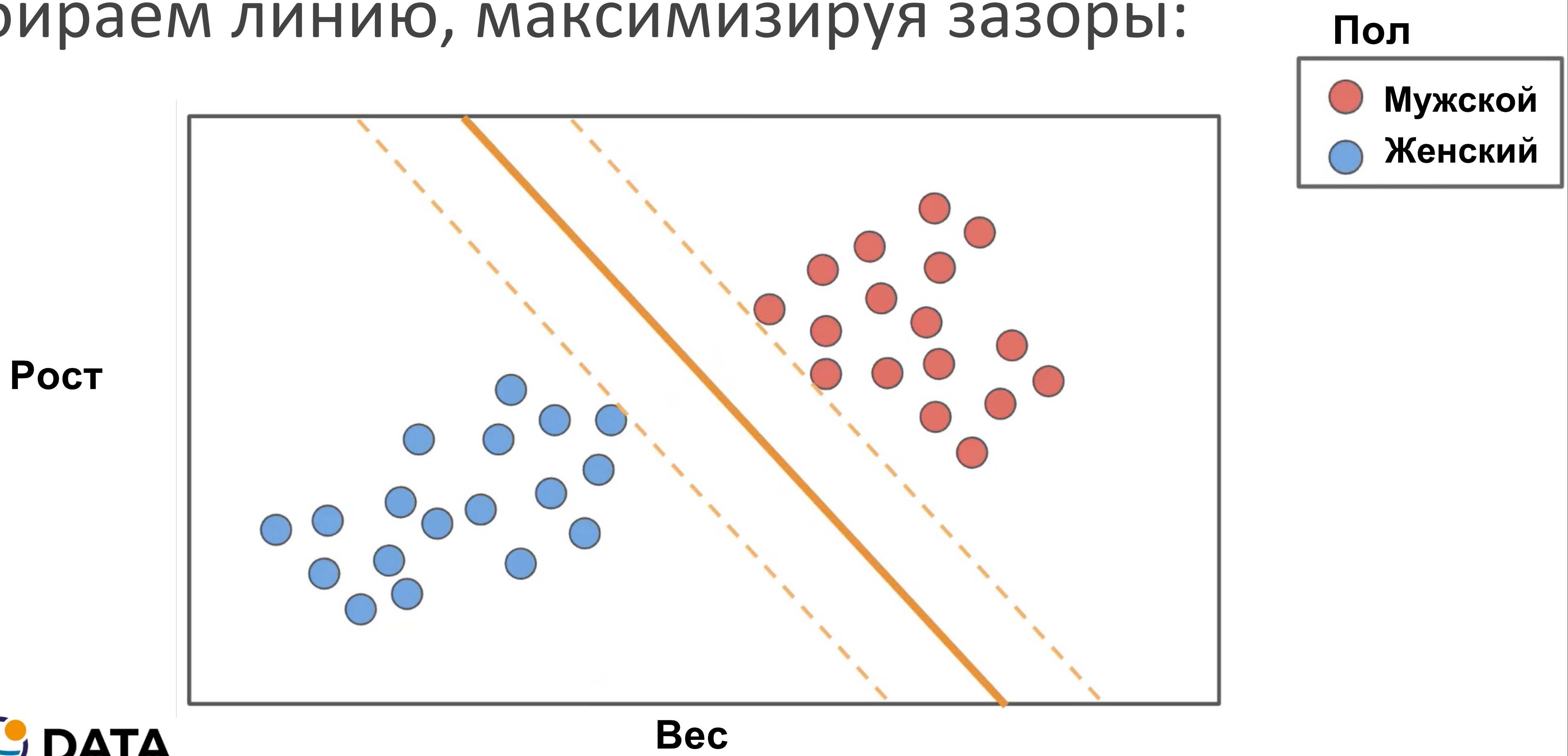
- Таких линий может быть множество:



PIERIAN DATA

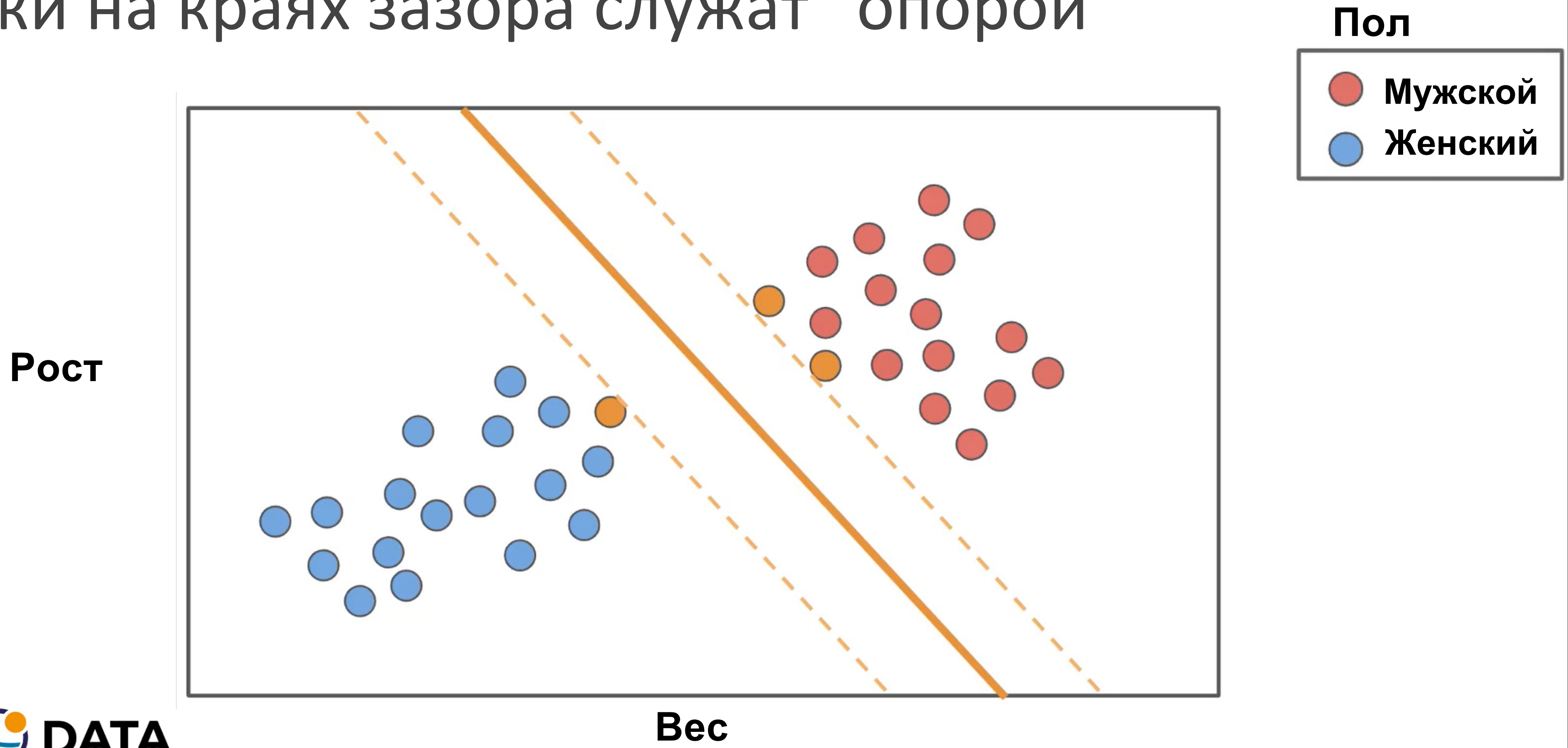
Метод опорных векторов

- Выбираем линию, максимизируя зазоры:



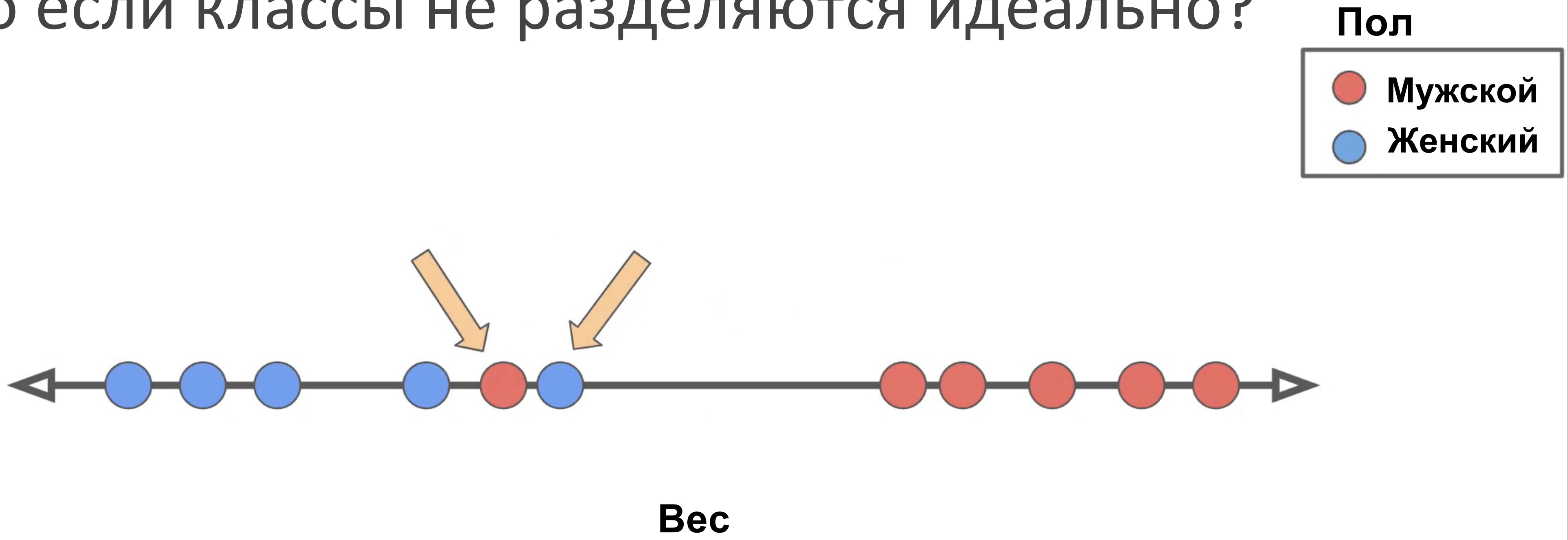
Метод опорных векторов

- Точки на краях зазора служат “опорой”



Метод опорных векторов

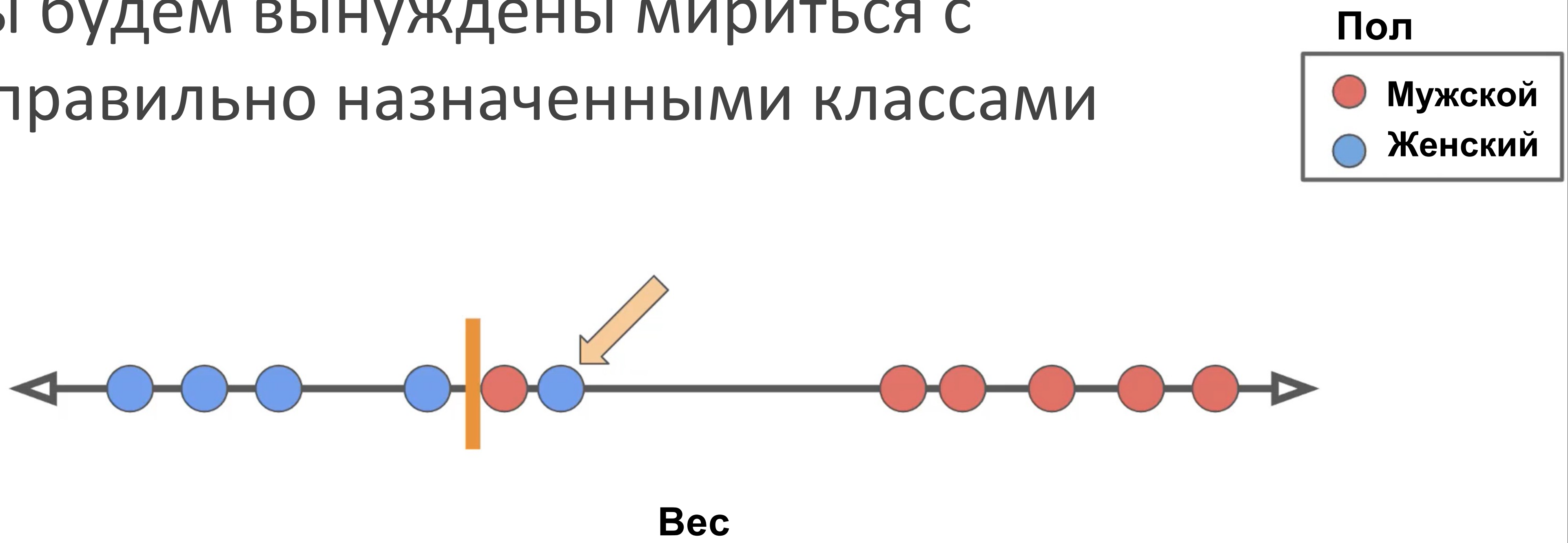
- Что если классы не разделяются идеально?



PIERIAN DATA

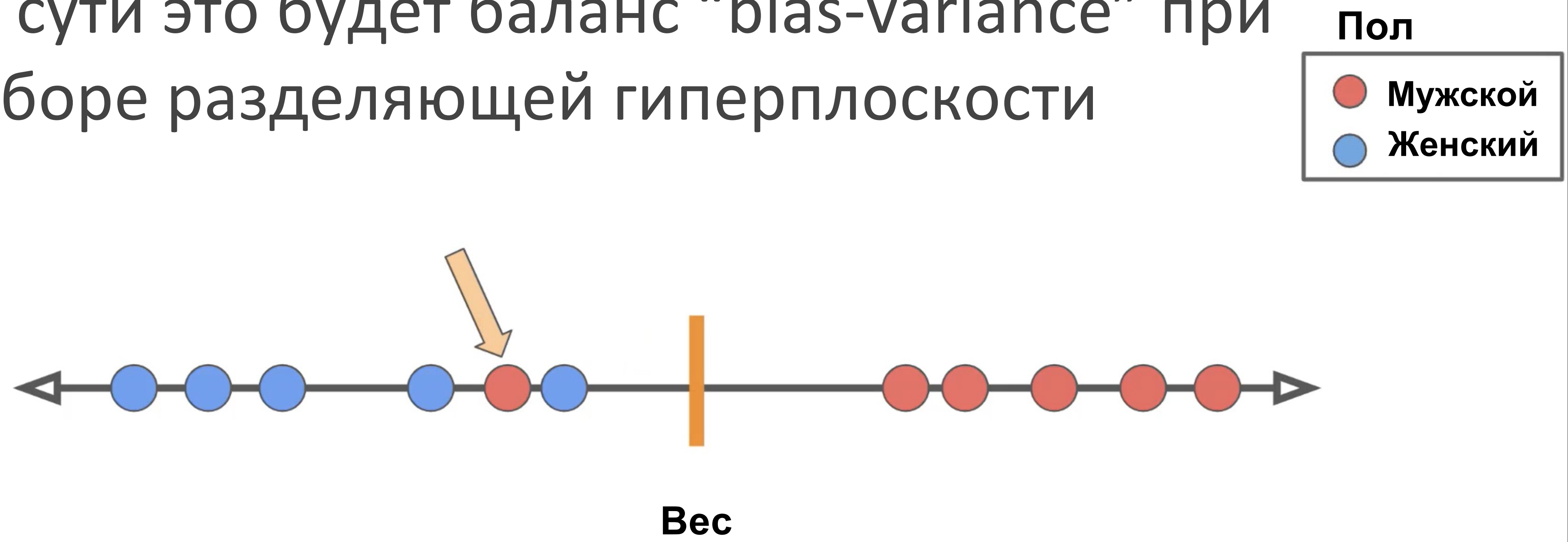
Метод опорных векторов

- Мы будем вынуждены мириться с неправильно назначенными классами



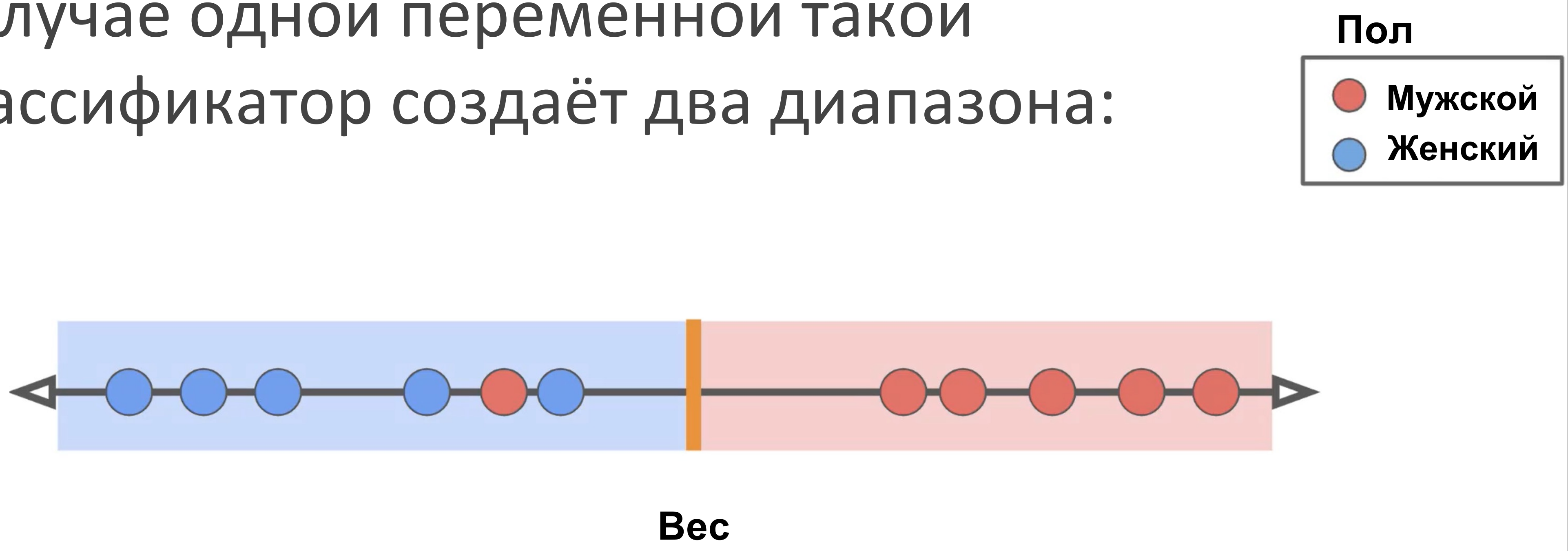
Метод опорных векторов

- По сути это будет баланс “bias-variance” при выборе разделяющей гиперплоскости



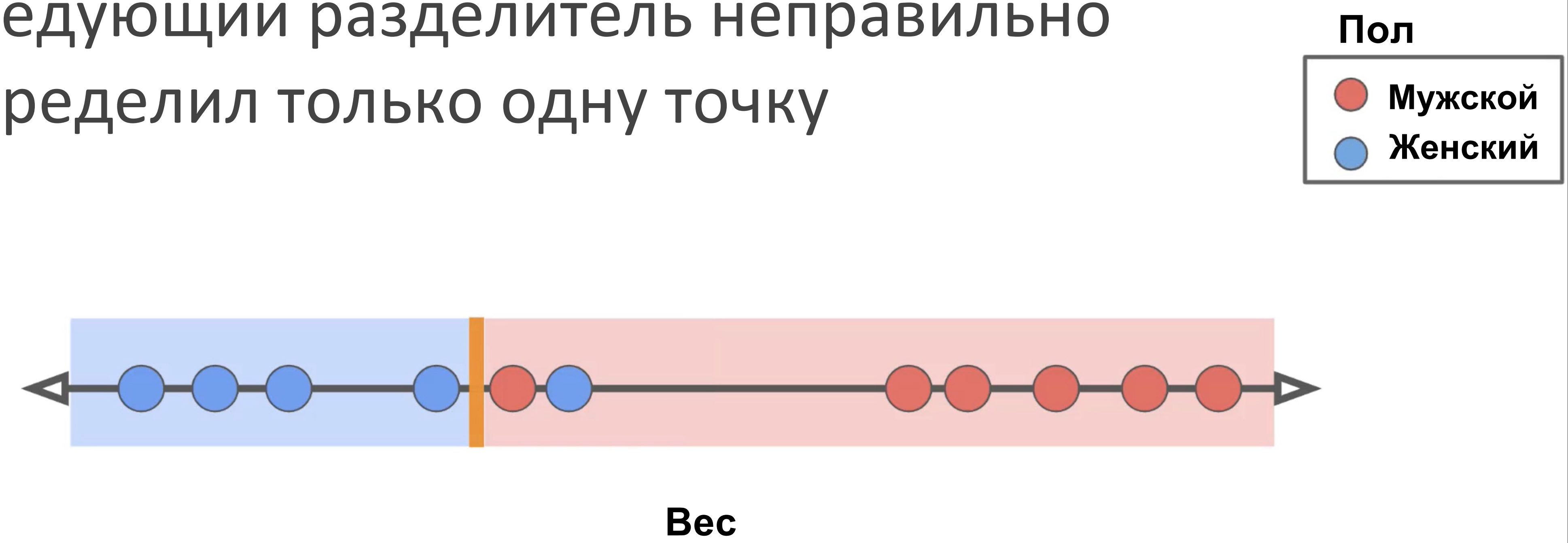
Метод опорных векторов

- В случае одной переменной такой классификатор создаёт два диапазона:



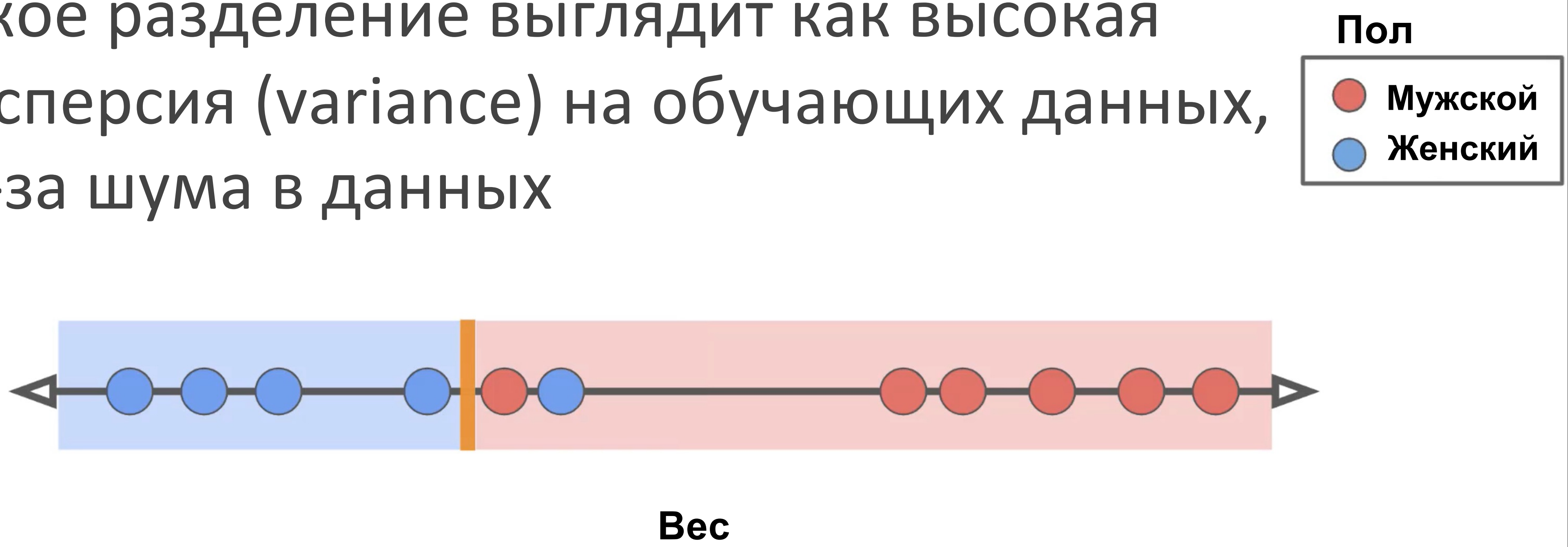
Метод опорных векторов

- Следующий разделитель неправильно определил только одну точку



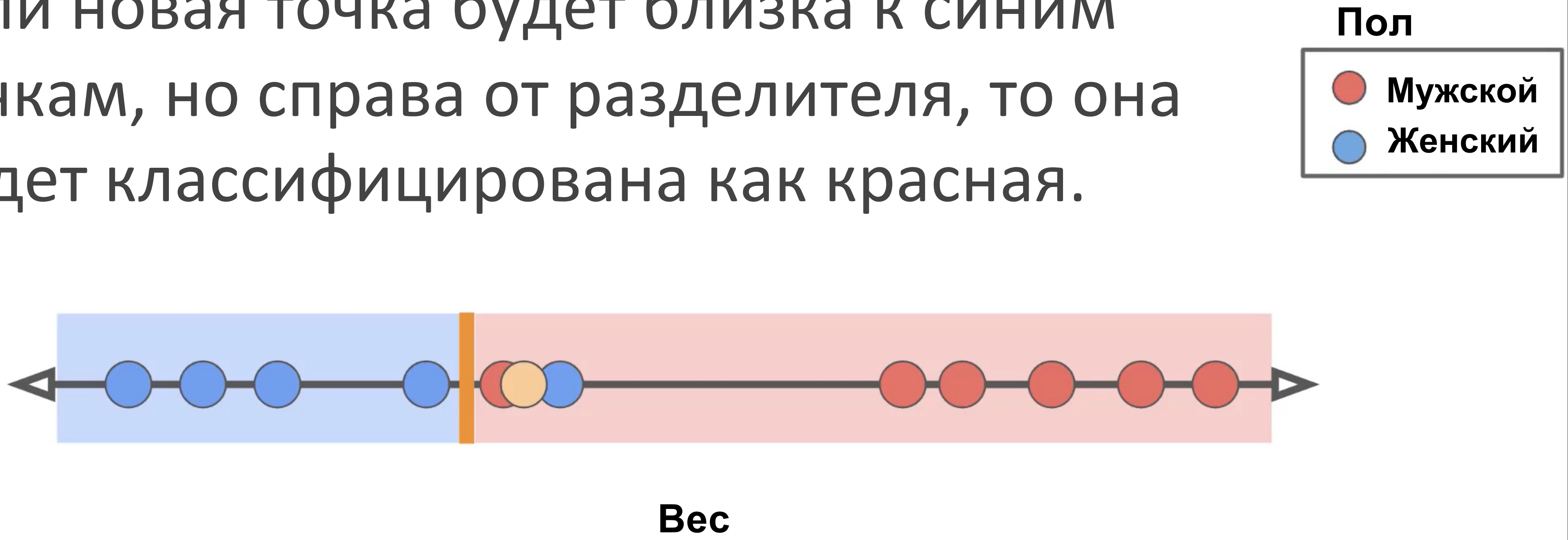
Метод опорных векторов

- Такое разделение выглядит как высокая дисперсия (variance) на обучающих данных, из-за шума в данных



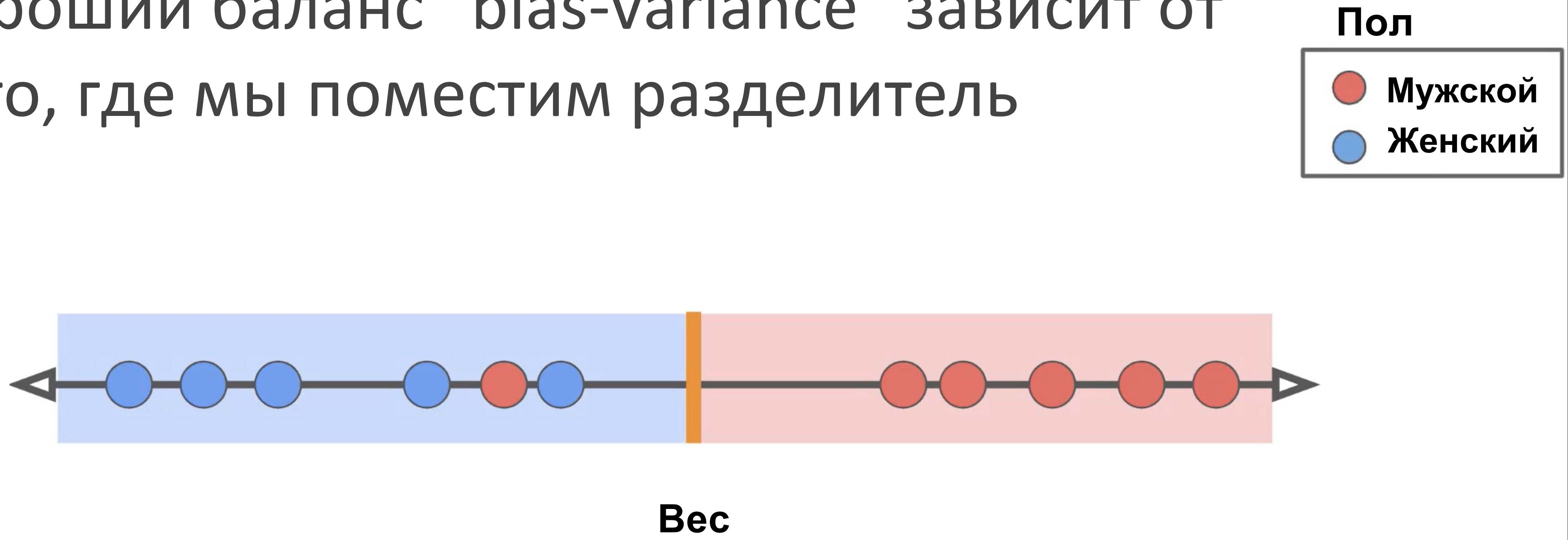
Метод опорных векторов

- Если новая точка будет близка к синим точкам, но справа от разделителя, то она будет классифицирована как красная.



Метод опорных векторов

- Хороший баланс “bias-variance” зависит от того, где мы поместим разделитель



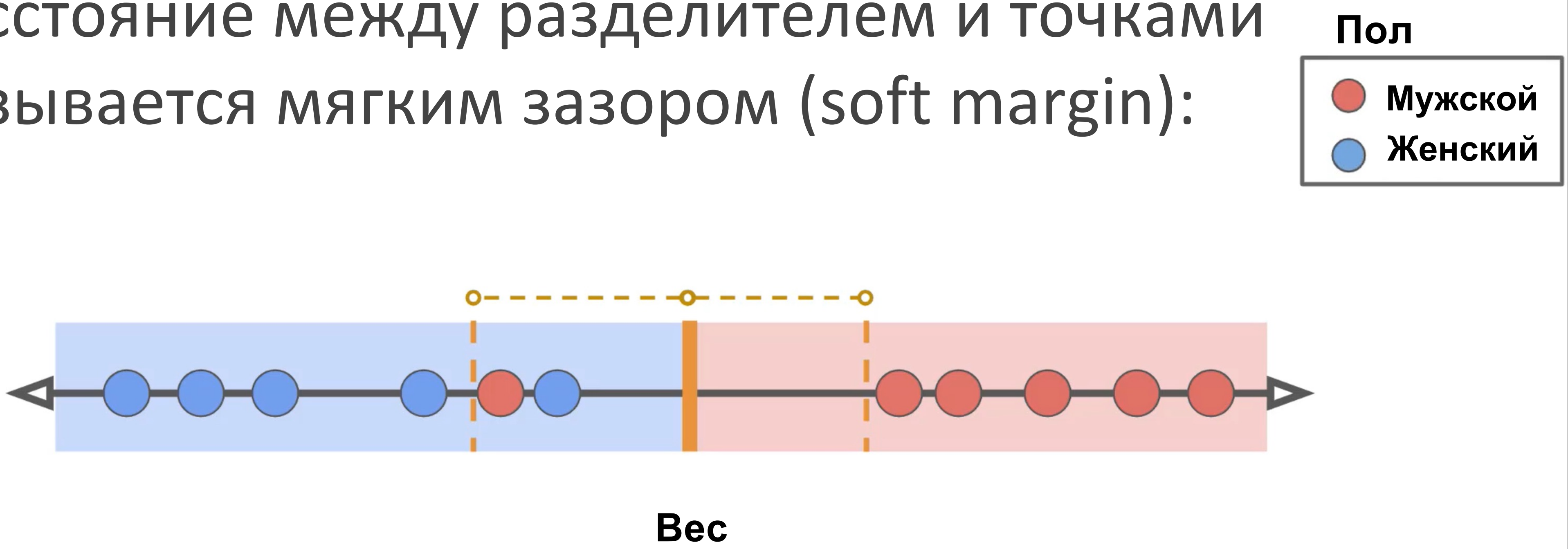
PIERIAN DATA

Метод опорных векторов

- Здесь мы допускаем большее смещение (bias), но модель лучше работает в будущем

Метод опорных векторов

- Расстояние между разделителем и точками называется мягким зазором (soft margin):

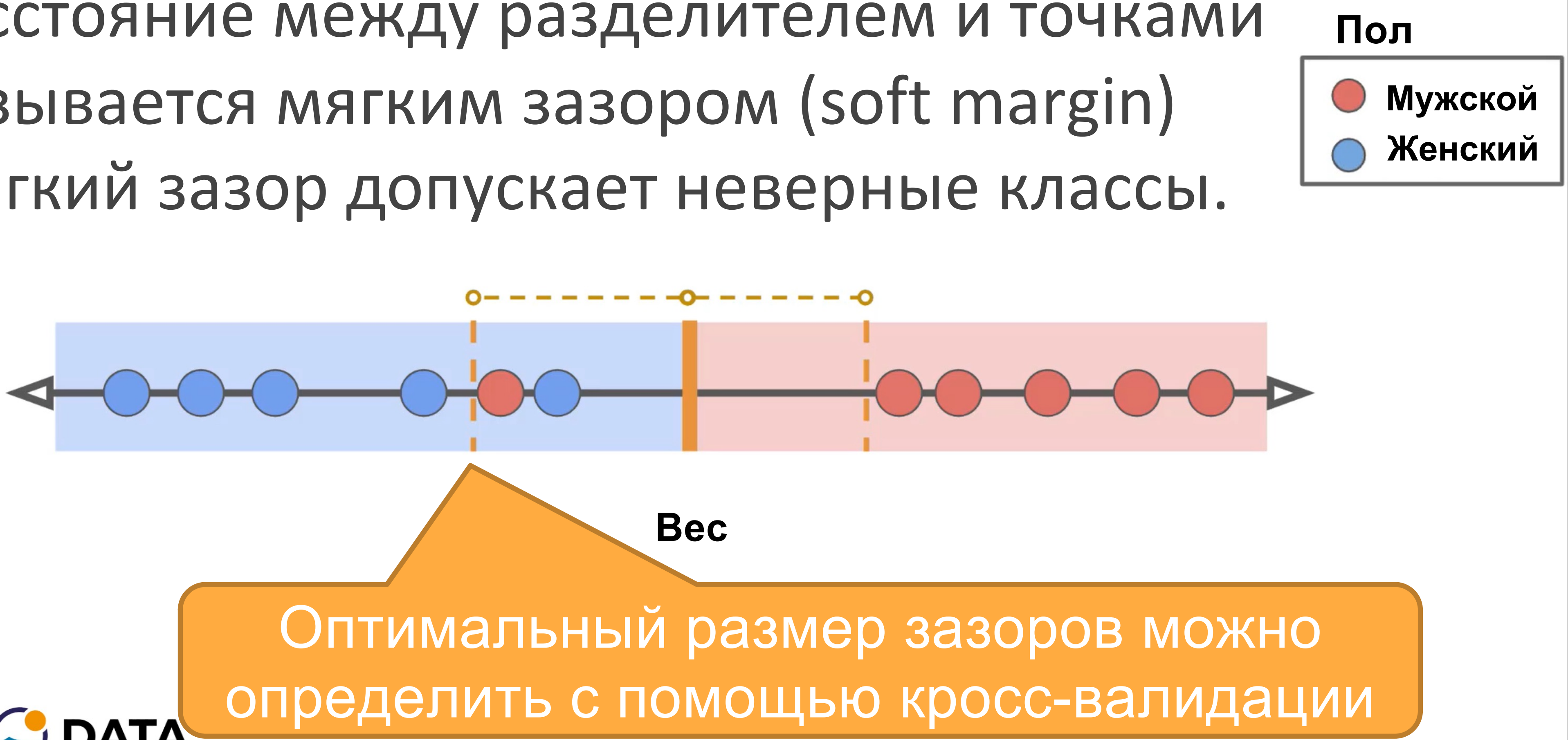


Метод опорных векторов

- Расстояние между разделителем и точками называется мягким зазором (soft margin)
- Мягкий зазор допускает неверные классы.

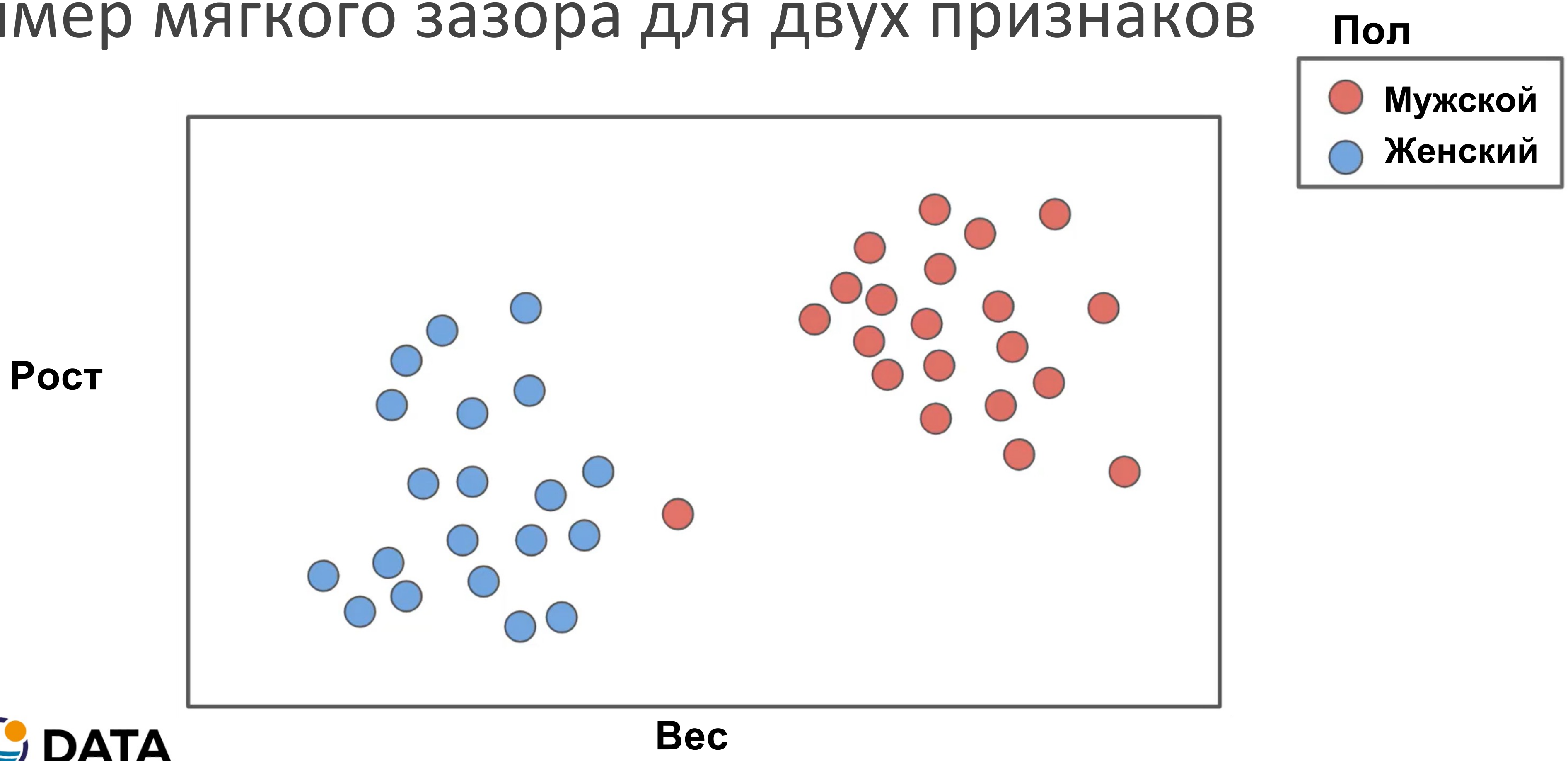
Метод опорных векторов

- Расстояние между разделителем и точками называется мягким зазором (soft margin)
- Мягкий зазор допускает неверные классы.



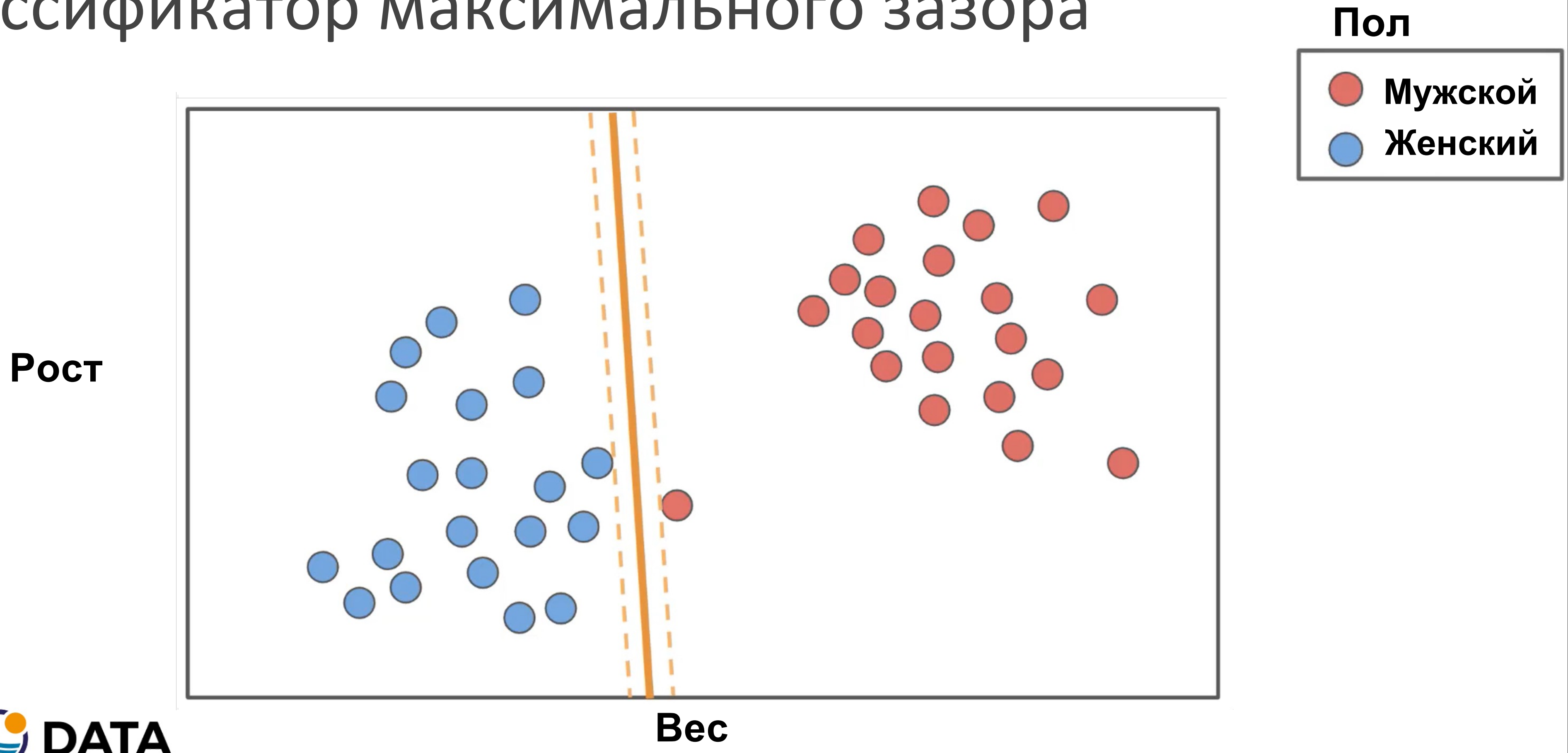
Метод опорных векторов

- Пример мягкого зазора для двух признаков



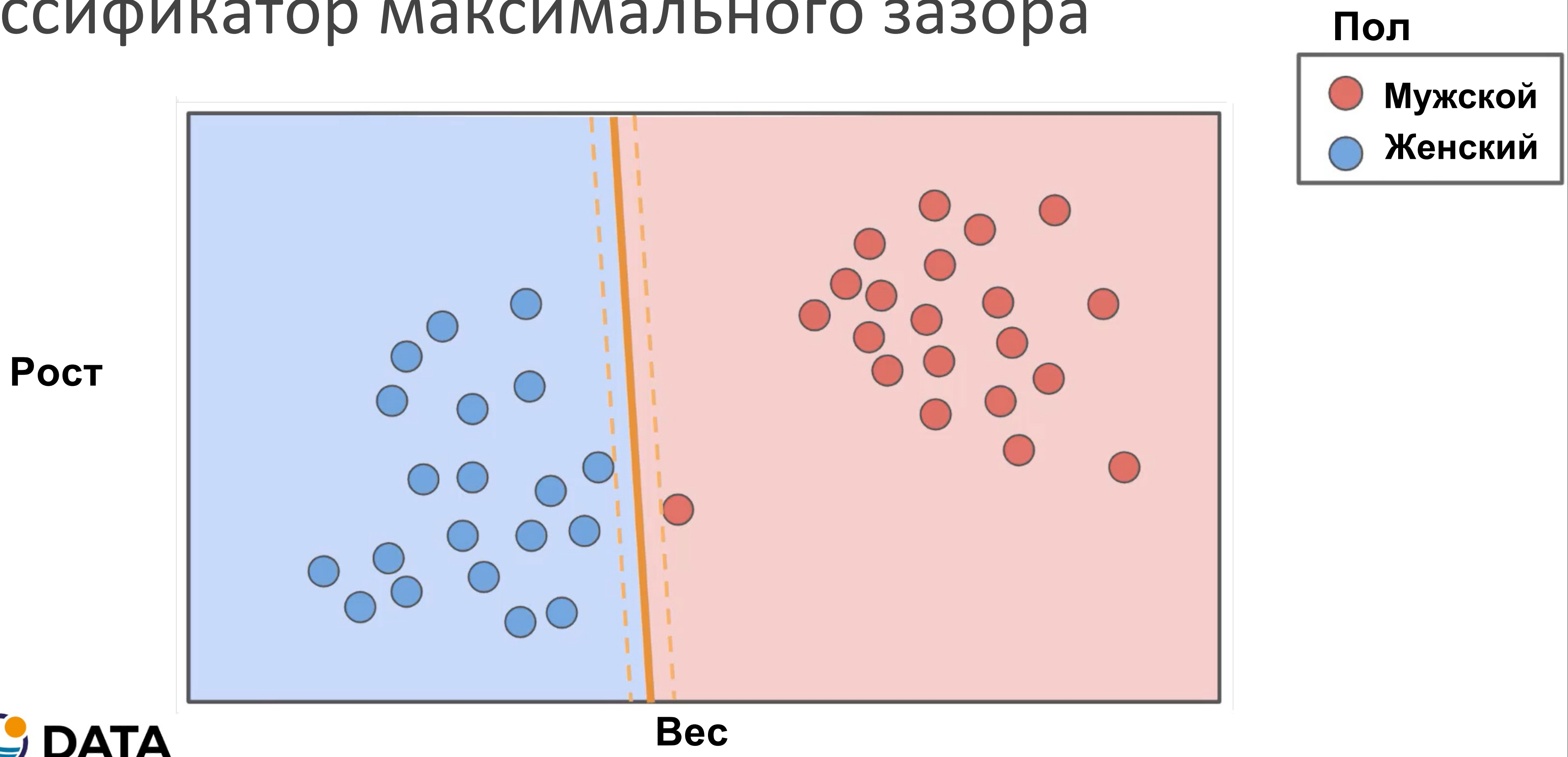
Метод опорных векторов

- Классификатор максимального зазора



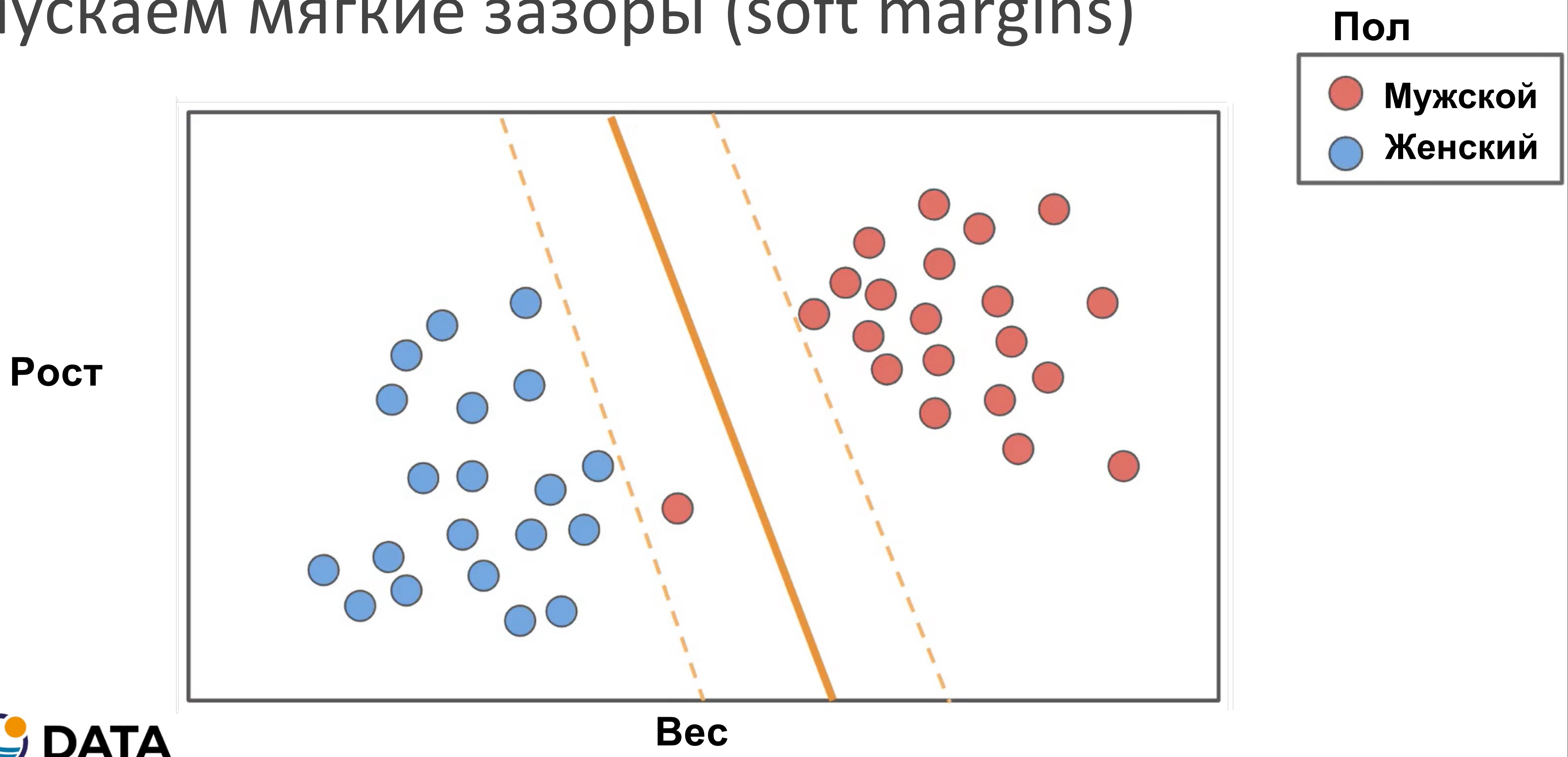
Метод опорных векторов

- Классификатор максимального зазора



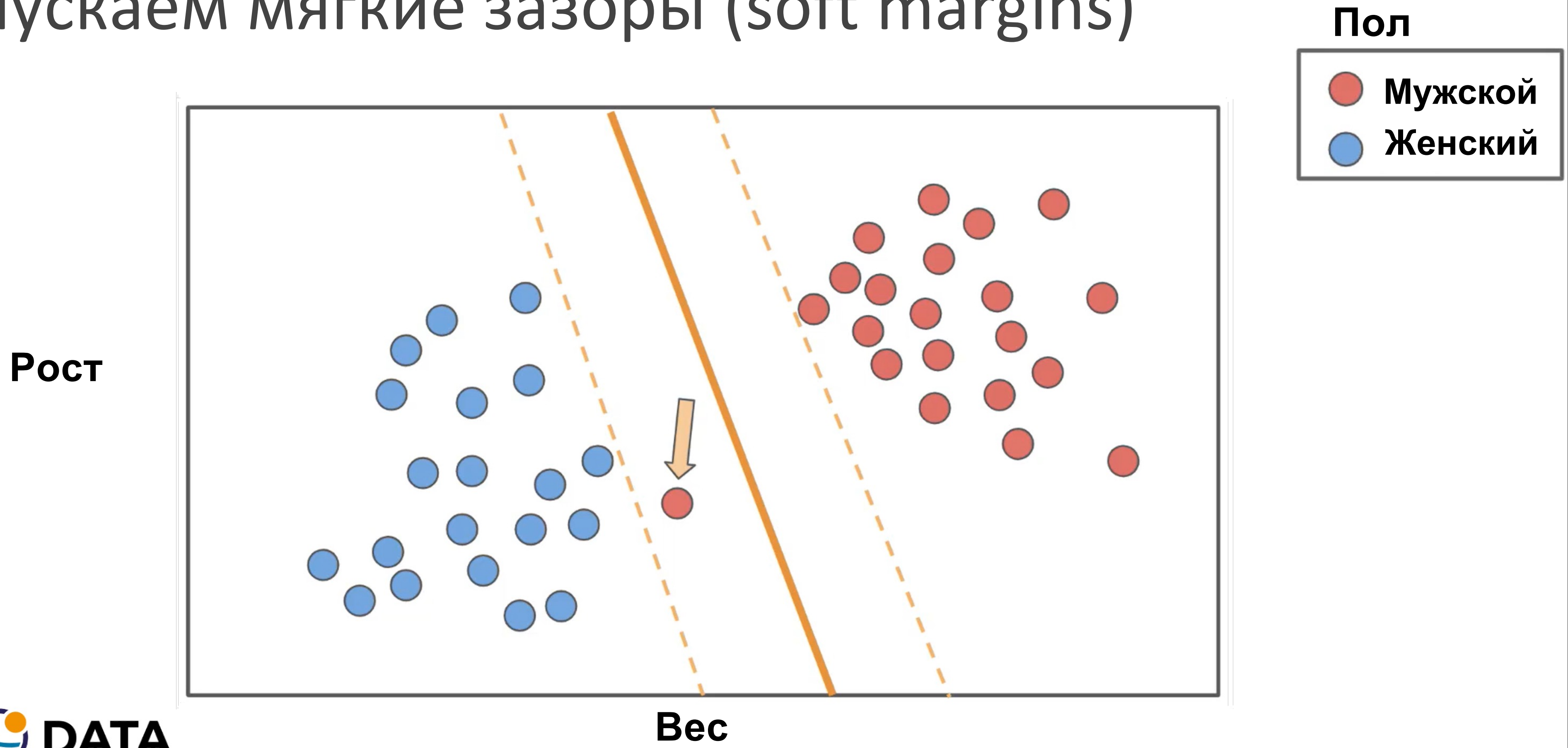
Метод опорных векторов

- Допускаем мягкие зазоры (soft margins)



Метод опорных векторов

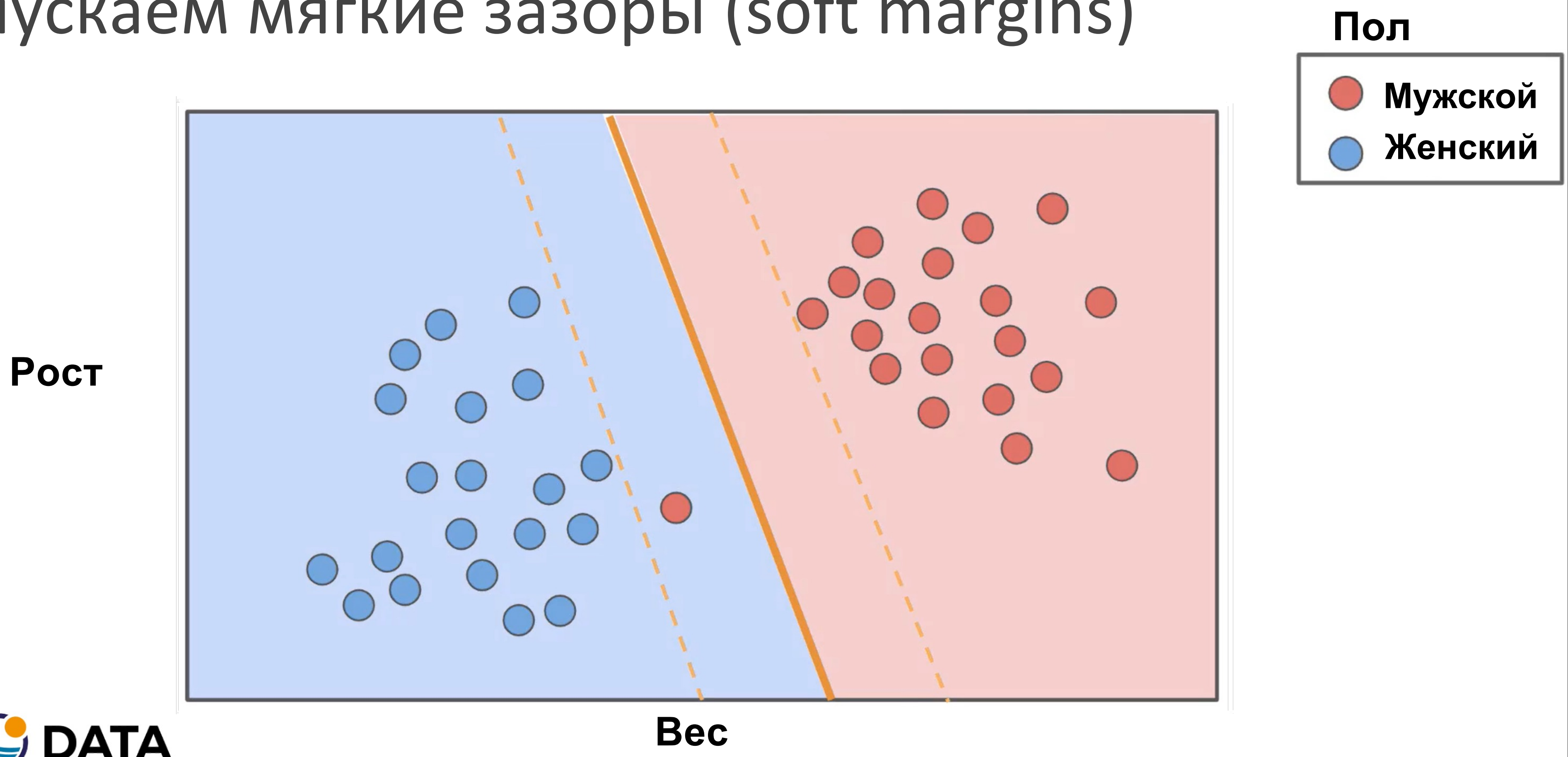
- Допускаем мягкие зазоры (soft margins)



PIERIAN DATA

Метод опорных векторов

- Допускаем мягкие зазоры (soft margins)

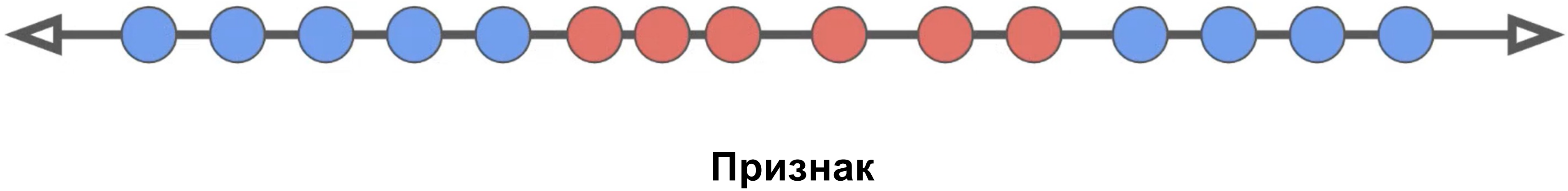


Метод опорных векторов

- Мы рассматривали только случаи, когда классы легко разделяются с помощью гиперплоскости.
- Допуская несколько неправильно определённых классов, мы в общем получили очень хорошие результаты.
- Но что если гиперплоскость показывает плохие результаты, даже если мы разрешаем неправильно определять классы?

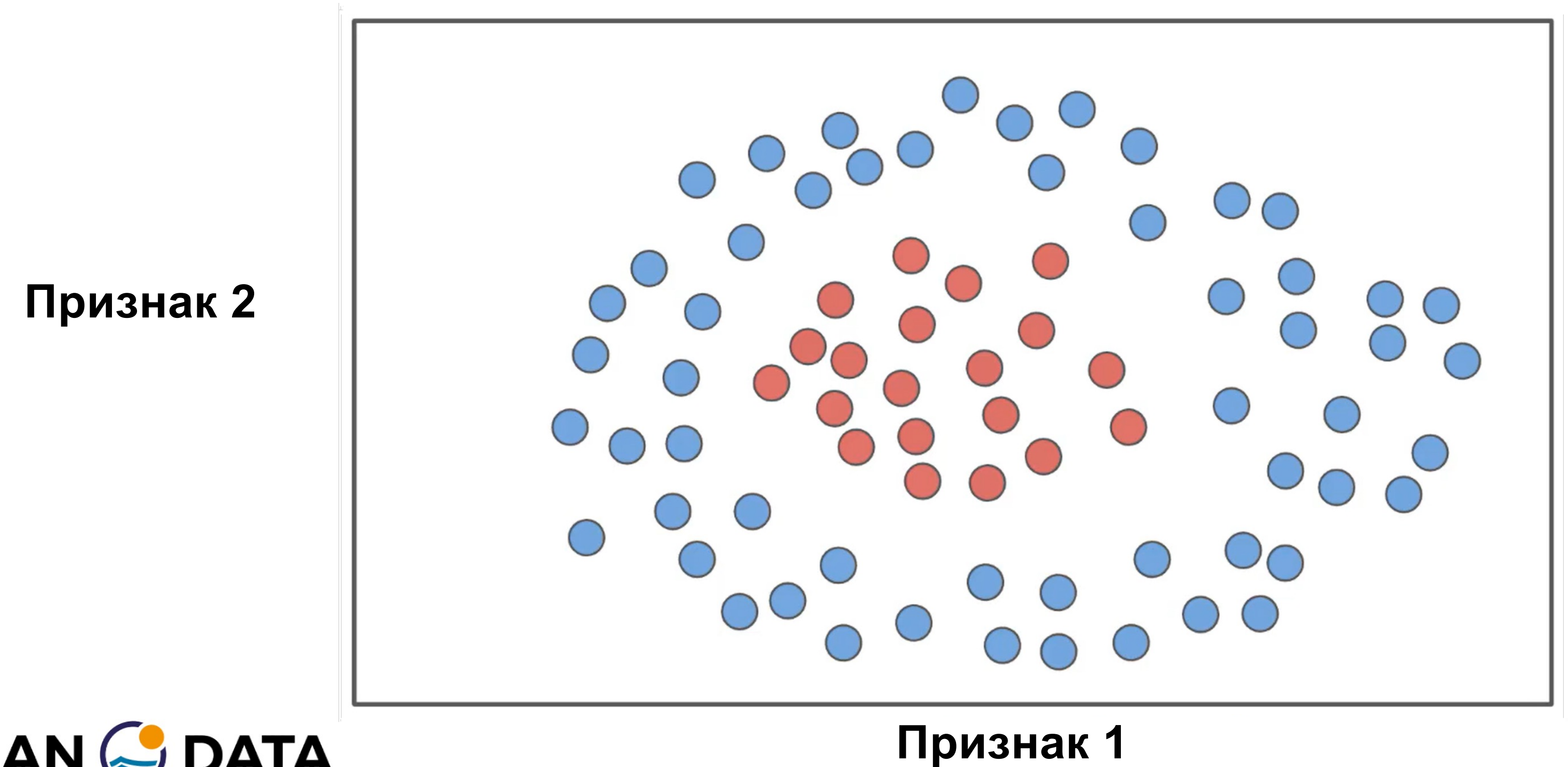
Метод опорных векторов

- В примере ниже одна гиперплоскость не может разбить классы – будет много неправильно классифицированных точек!



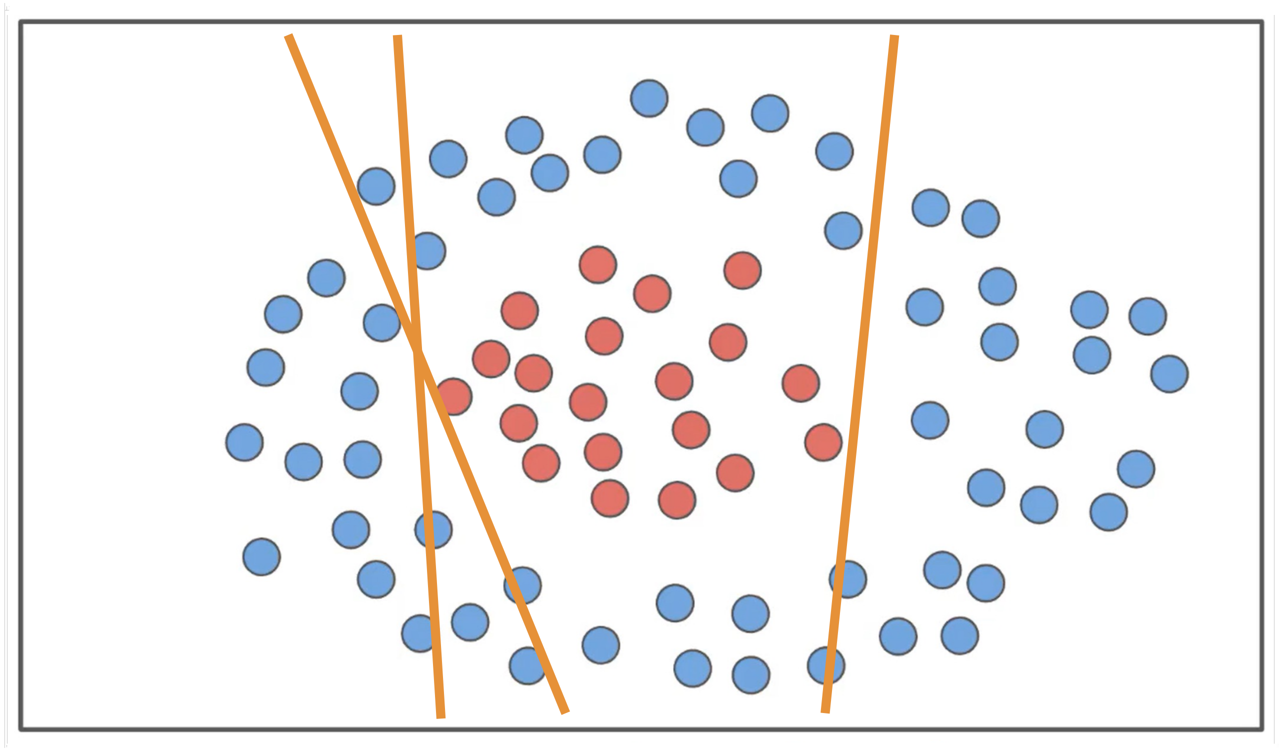
Метод опорных векторов

- Для двумерного пространства:



Метод опорных векторов

- Для двумерного пространства:



PIERIAN DATA

Метод опорных векторов

- Чтобы решить задачу для таких случаев, в методе опорных векторов используются ядра (kernels).
- С их помощью данные проецируются в пространство большей размерности, и уже в нём применяется гиперплоскость для разделения данных на классы.

Метод опорных векторов (support vector machines)

Теория – ядра

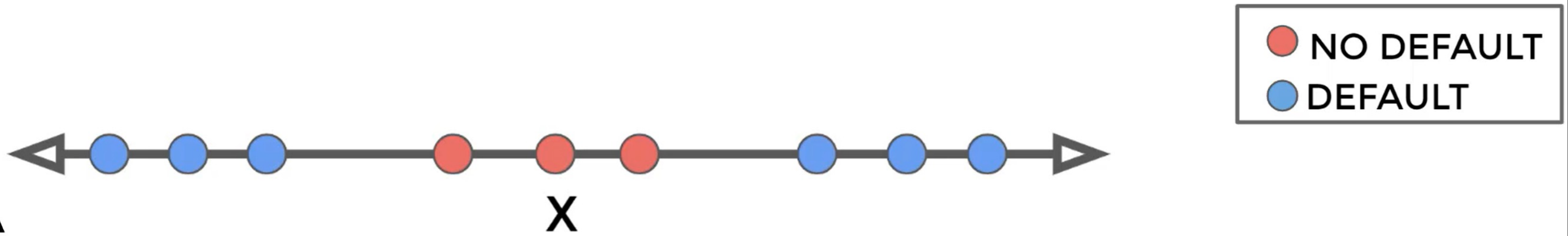
PIERIAN DATA

Метод опорных векторов

- Ядра (kernels) позволят нам перейти от классификатора опорных векторов к полноценному методу опорных векторов.
- Существует множество разных ядер, с помощью которых можно спроецировать данные в пространство большей размерности.
- Посмотрим несколько визуальных примеров!

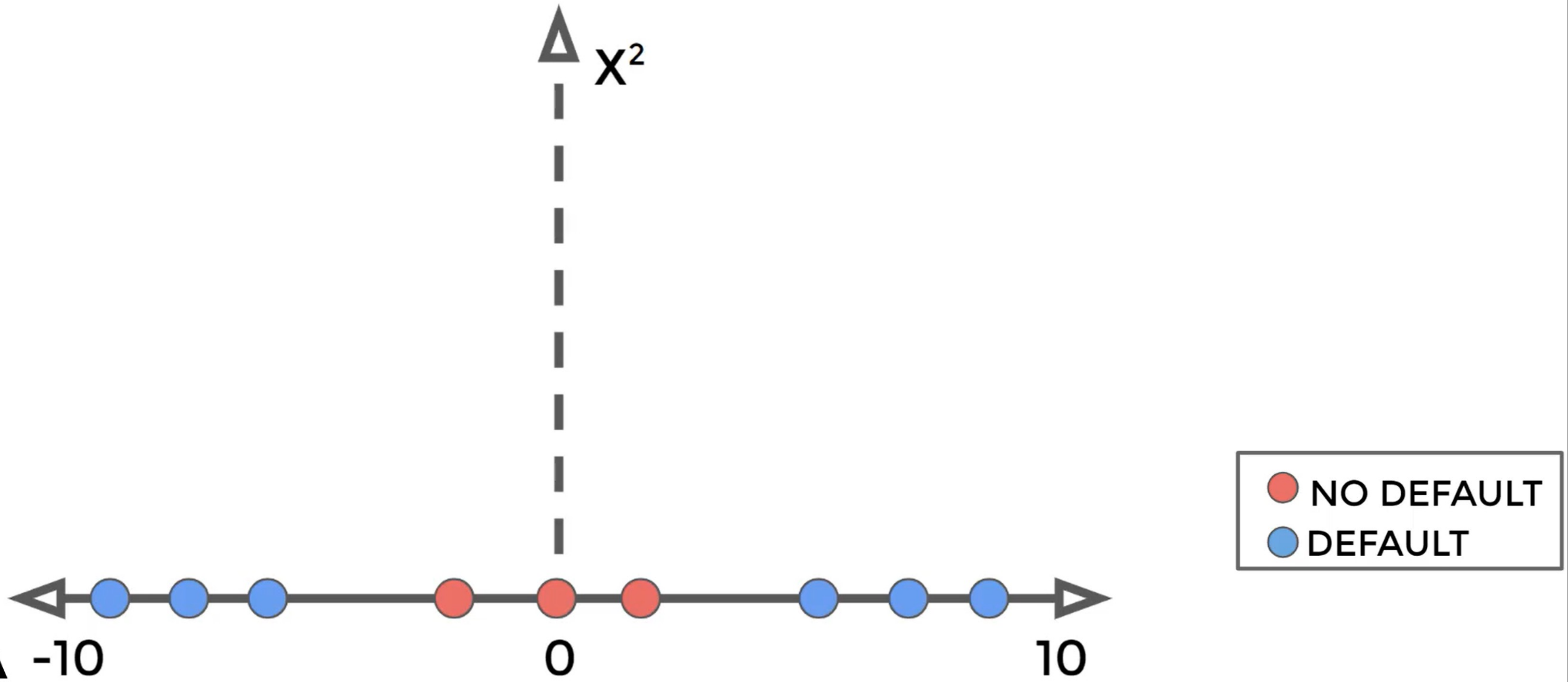
Метод опорных векторов

- Вспомним наш 1-мерный пример, в котором не получалось разделить данные гиперплоскостью



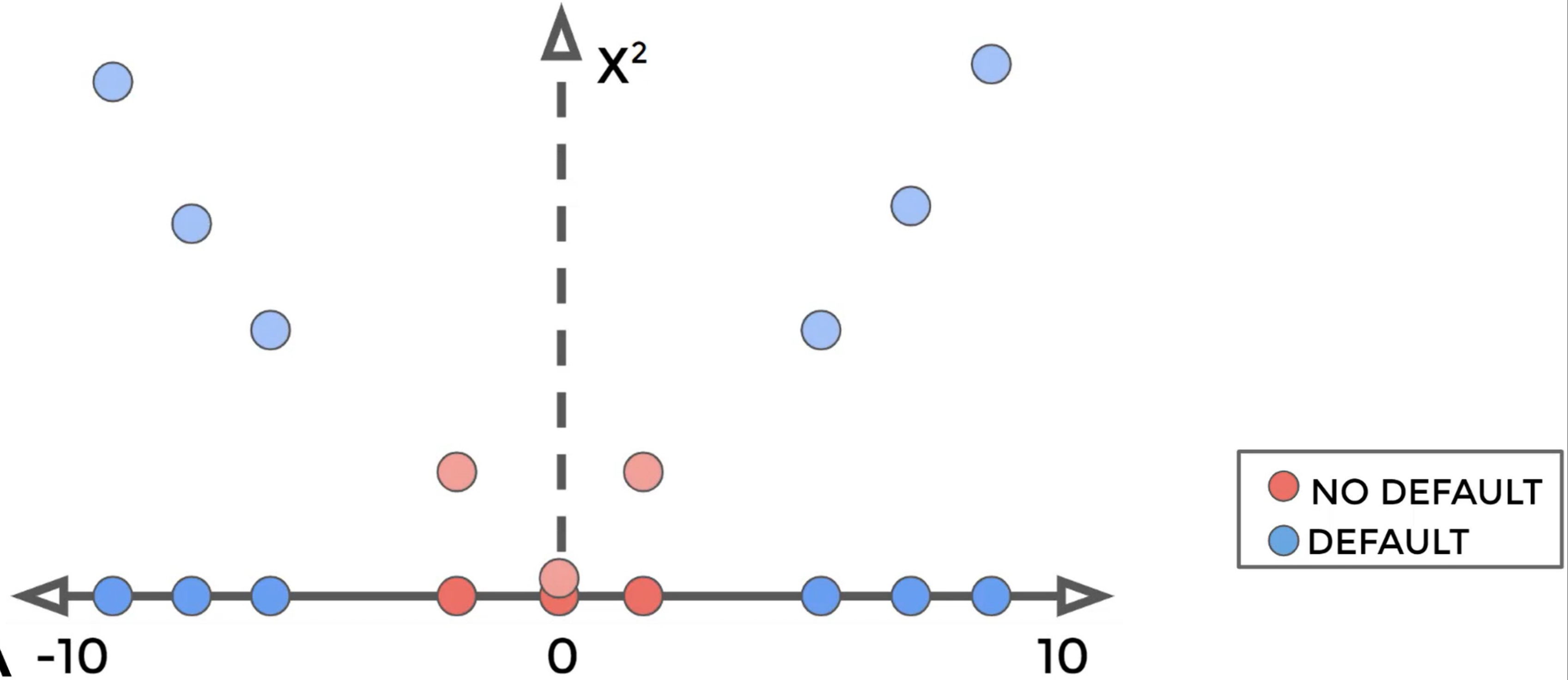
Метод опорных векторов

- Попробуем применить полиномиальное ядро, чтобы добавить измерение X^2



Метод опорных векторов

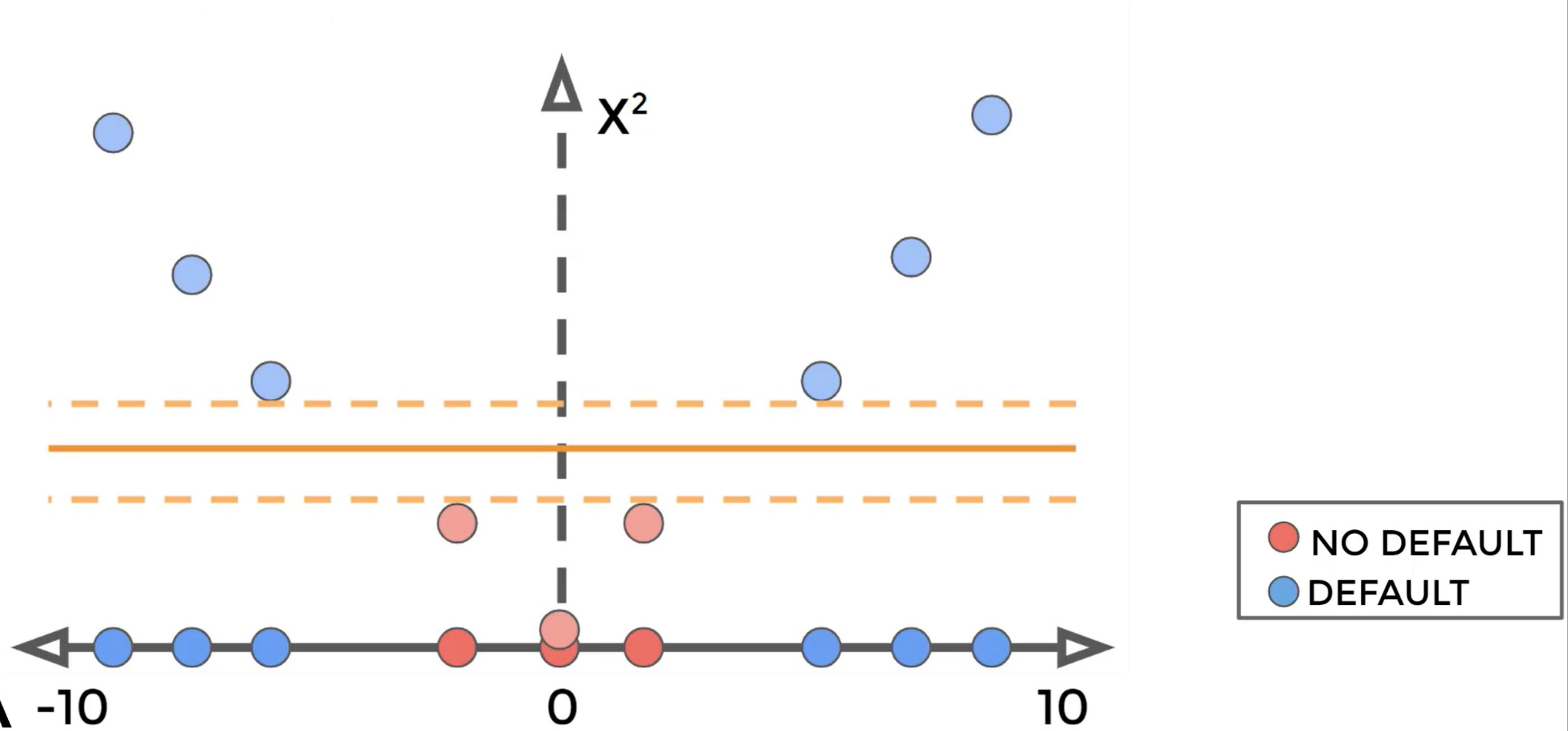
- Попробуем применить полиномиальное ядро, чтобы добавить измерение X^2



PIERIAN DATA

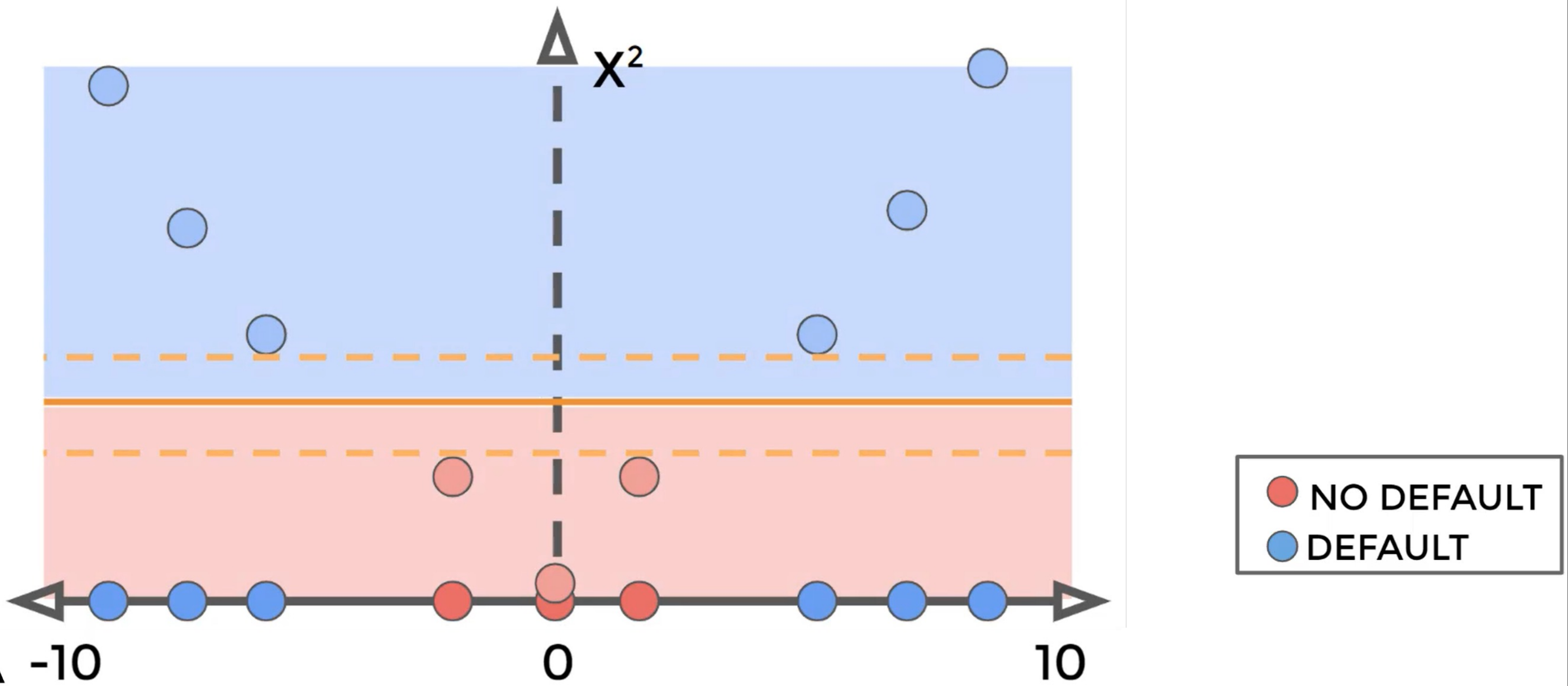
Метод опорных векторов

- После проекции строим гиперплоскость:



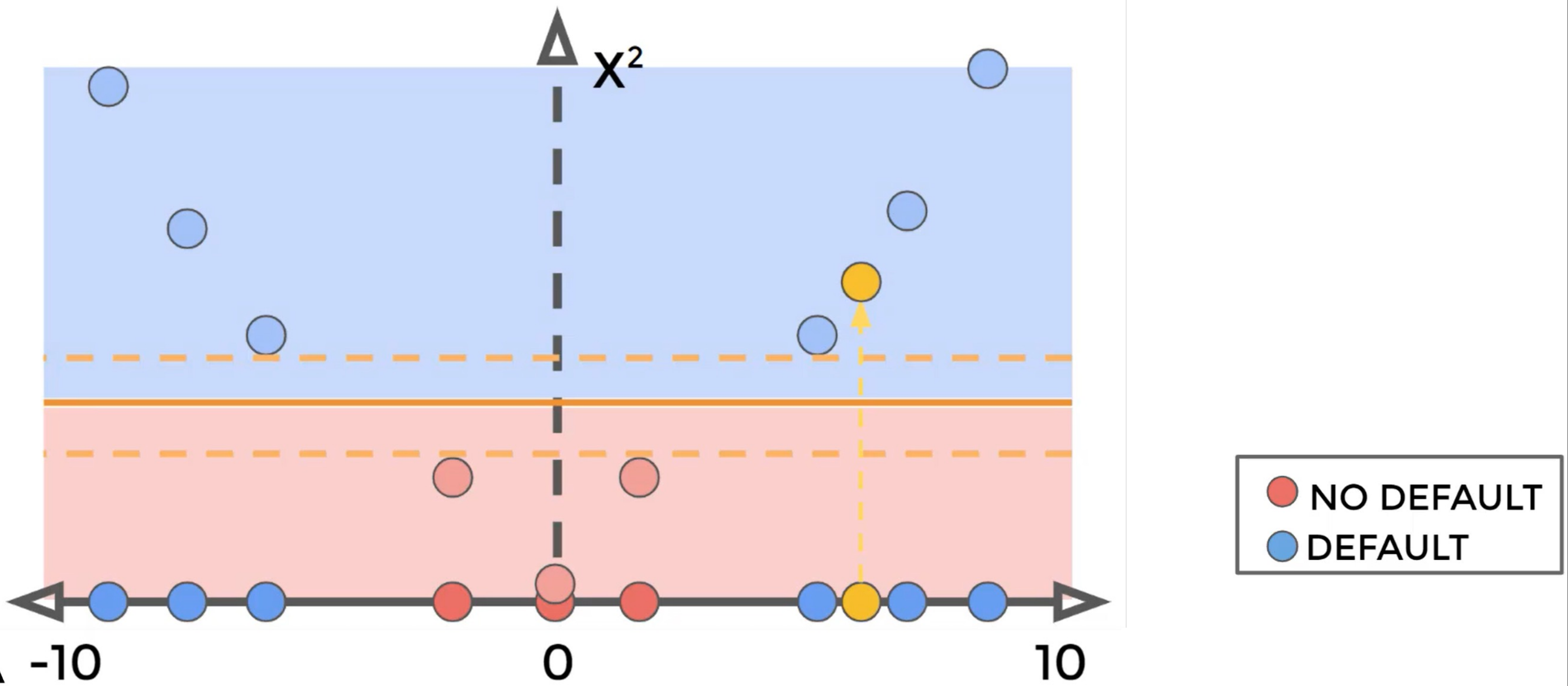
Метод опорных векторов

- После проекции строим гиперплоскость:



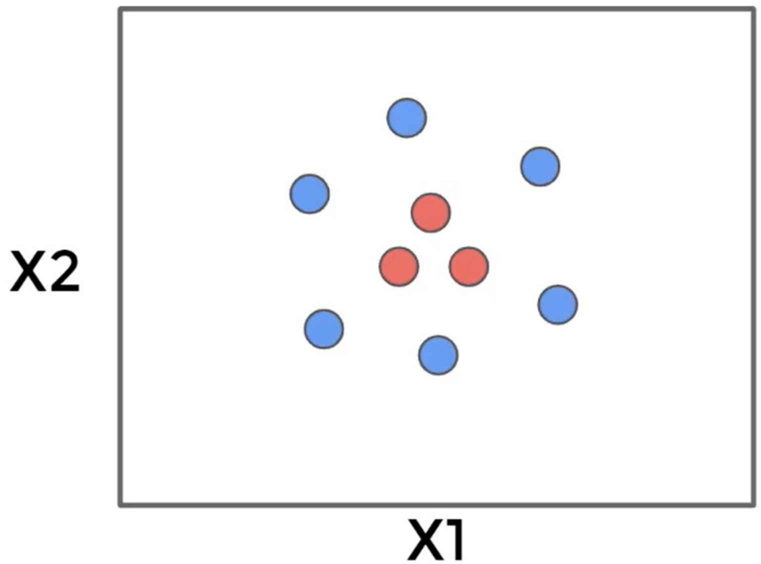
Метод опорных векторов

- С помощью этой гиперплоскости классифицируем новые точки:



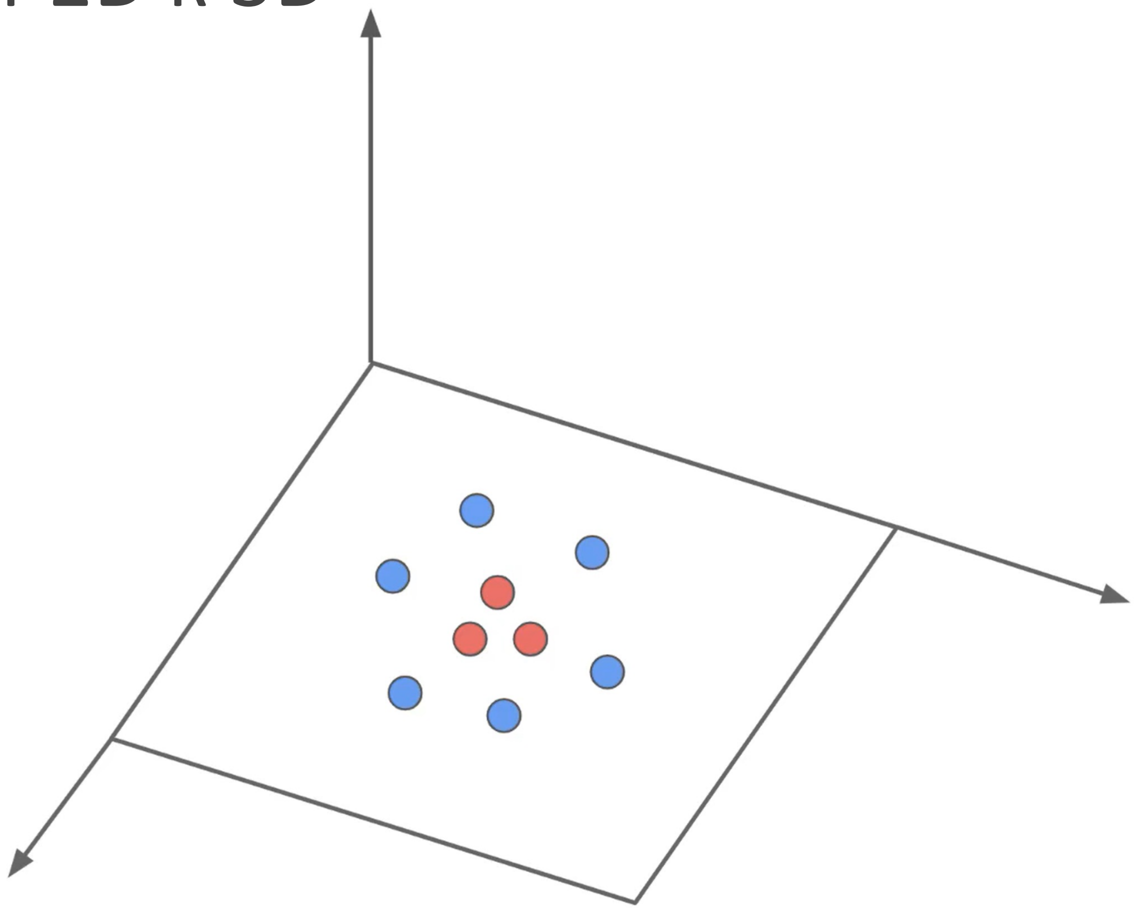
Метод опорных векторов

- Перейдём теперь к 2-мерному примеру, в котором не получалось разделить данные гиперплоскостью



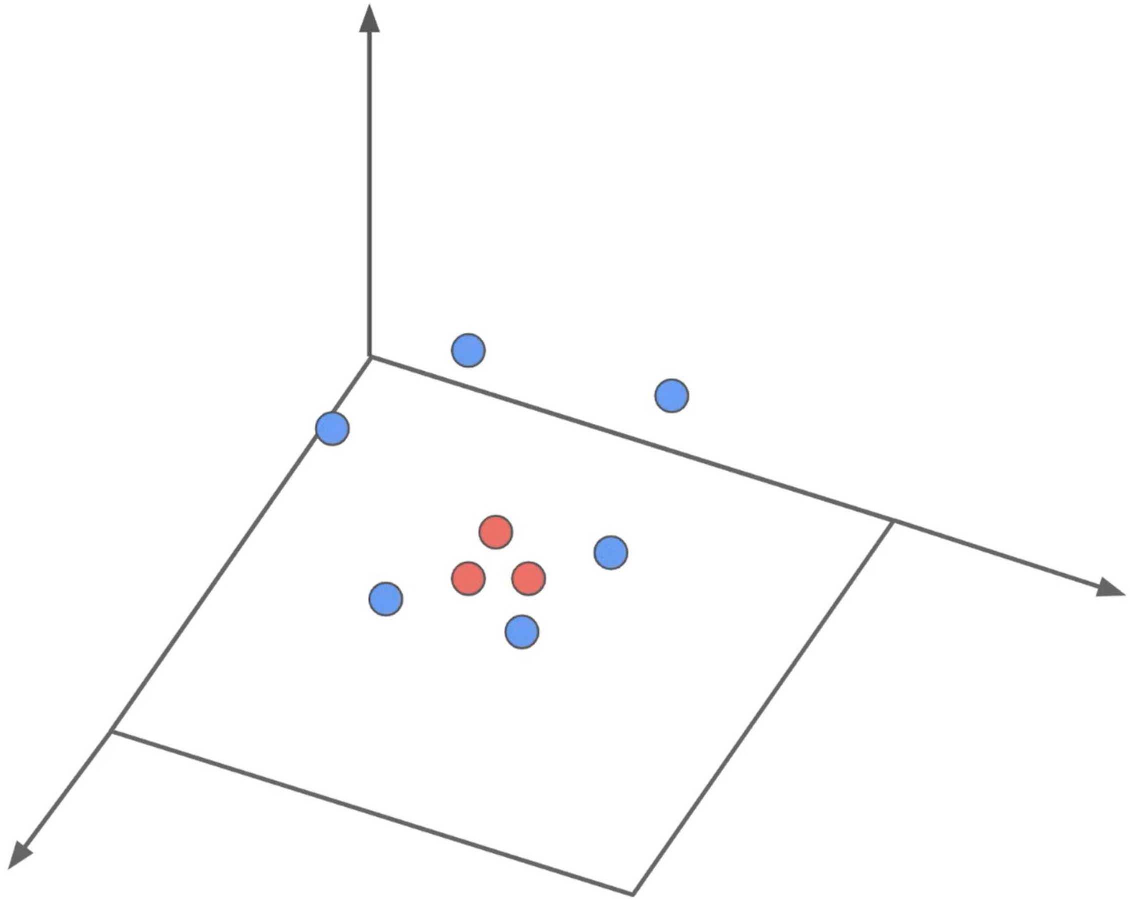
Метод опорных векторов

- Перейдём от 2D к 3D



Метод опорных векторов

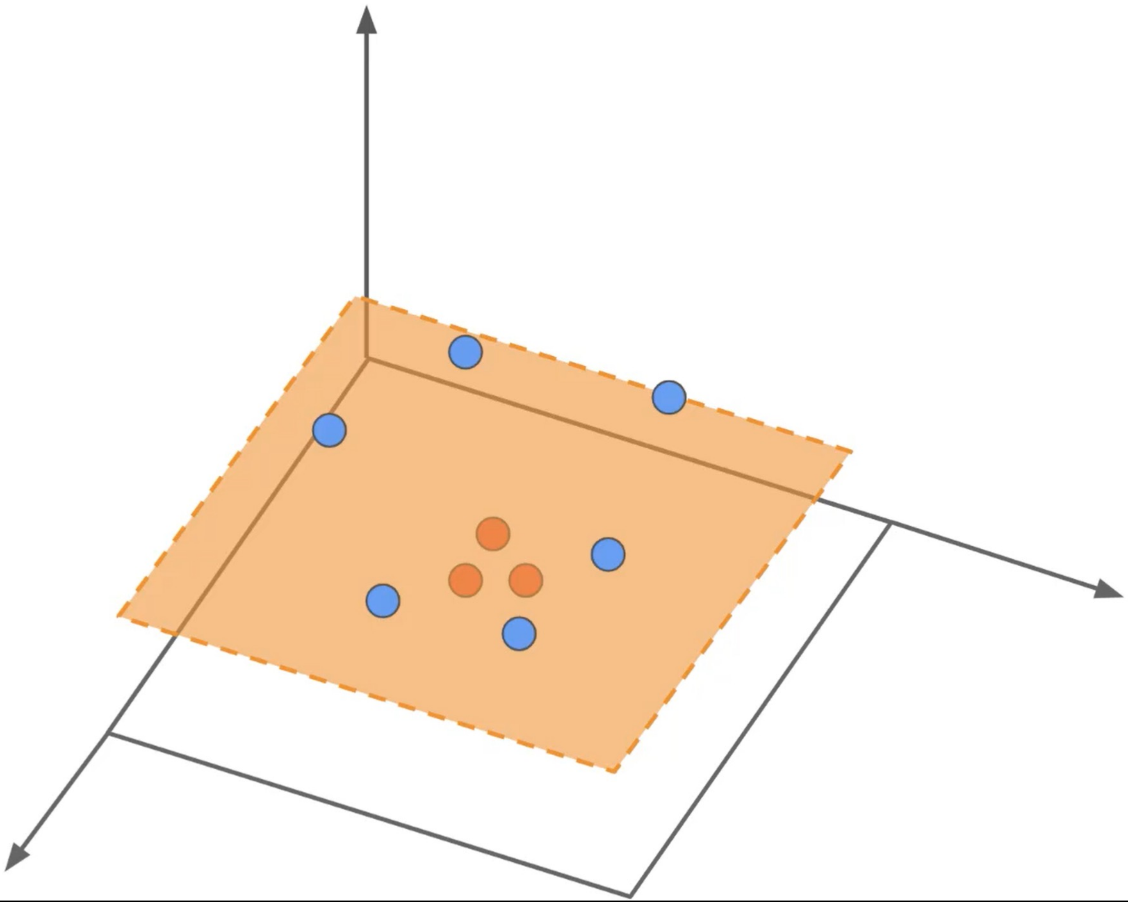
- Перейдём от 2D к 3D



PIERIAN DATA

Метод опорных векторов

- Гиперплоскость



Метод опорных векторов

- Возможно Вы слышали, что применение ядер в методе SVM ещё называют “kernel trick”.
- Мы переводим данные в пространство большей размерности.
- Математически, kernel trick позволяет не пересчитывать все точки в пространстве большей размерности.
- Почему? Узнаем в следующей лекции.

Метод опорных векторов (support vector machines)

Теория – “kernel trick” и математика

PIERIAN DATA

Метод опорных векторов

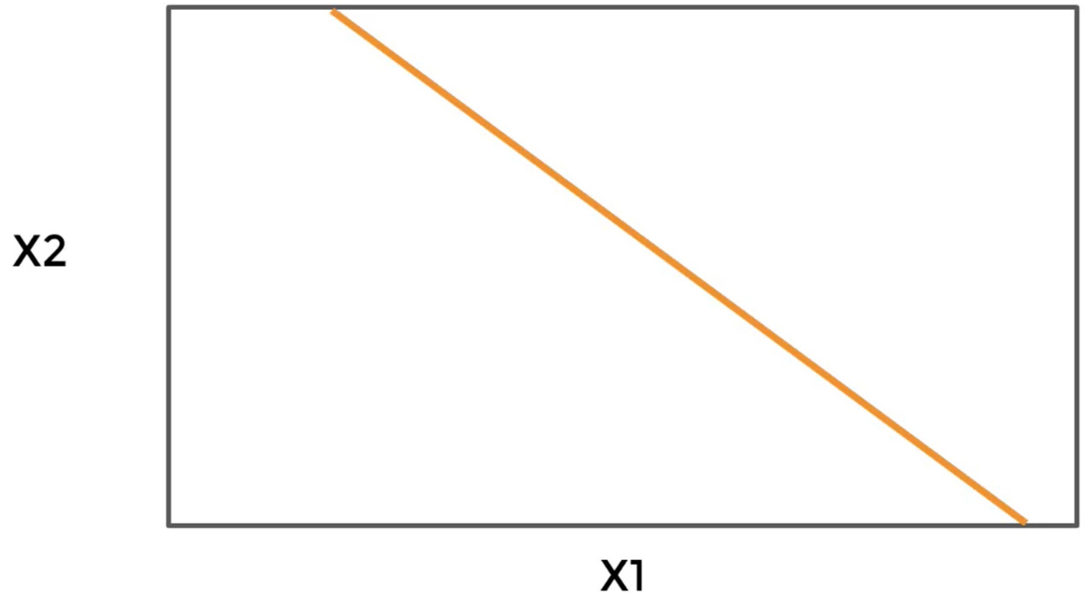
- Давайте посмотрим на общие принципы математики метода SVM, и как они соотносятся с командами Scikit-Learn.
- Мы начнём с обзора классификаторов с зазорами, и как их можно описать с помощью уравнений.

  - Замечание: можете рассматривать эту лекцию как “опциональную”.

Метод опорных векторов

- Определение гиперплоскости

$$\beta_0 + \beta_1 X_1 + \beta_2 X_2 = 0$$



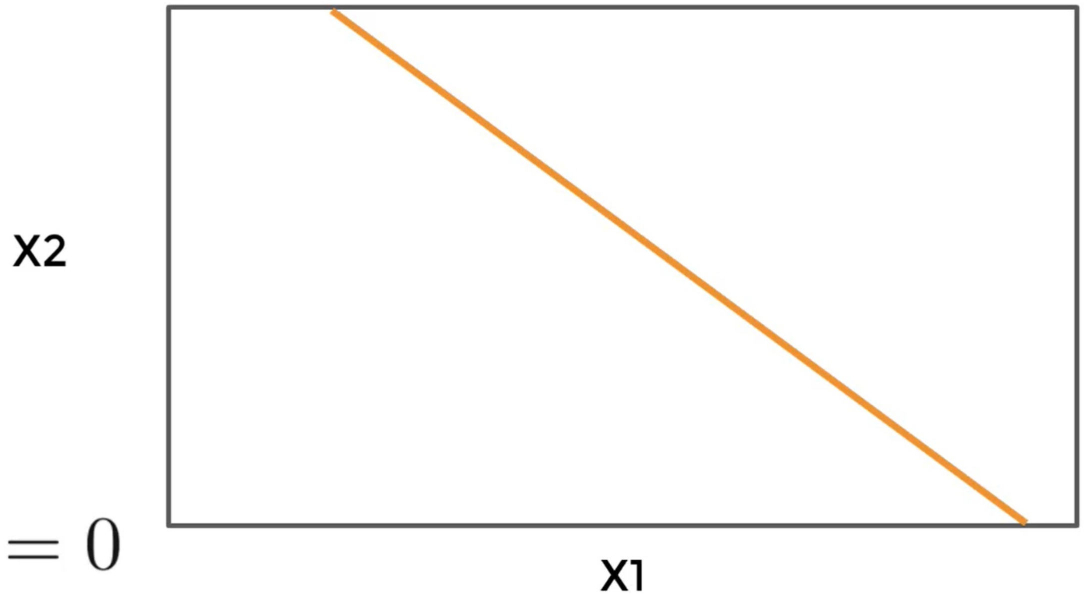
Метод опорных векторов

- Определение гиперплоскости

$$\beta_0 + \beta_1 X_1 + \beta_2 X_2 = 0$$

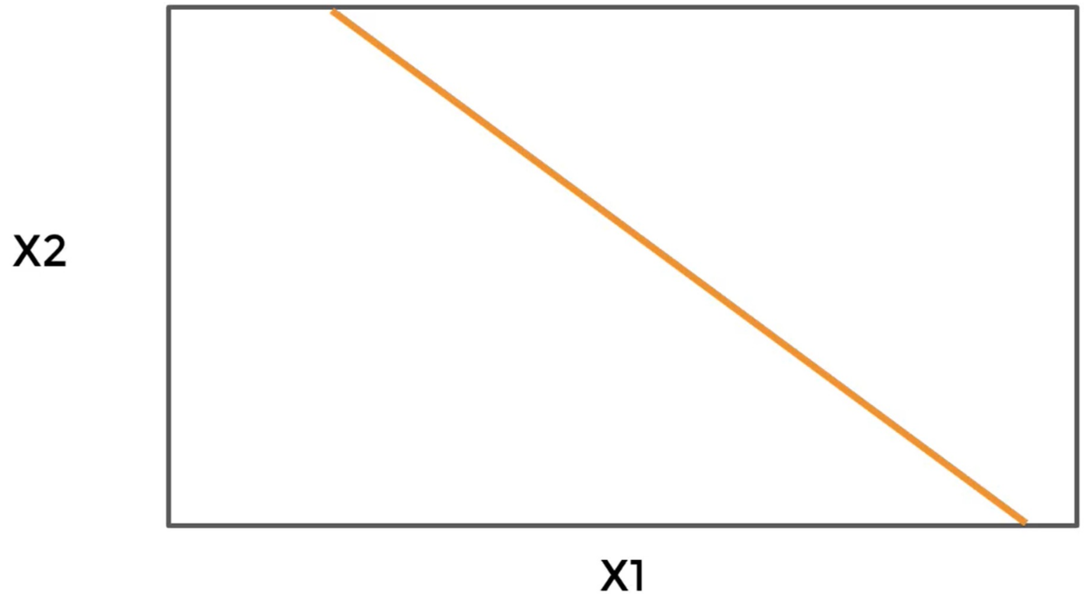
$$\Downarrow$$

$$\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p = 0$$



Метод опорных векторов

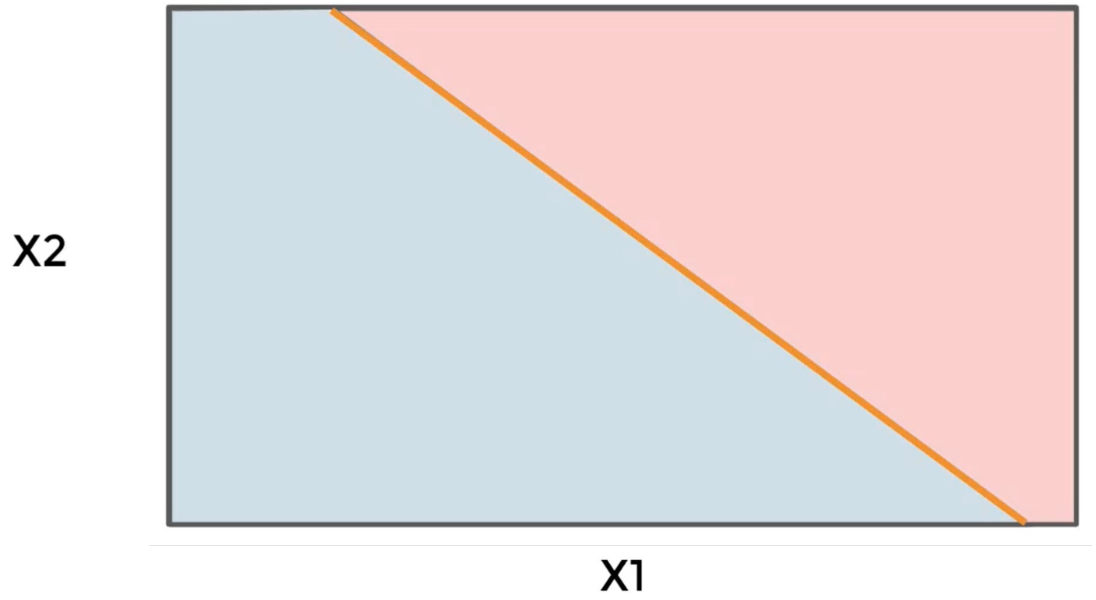
- Разделение пространства гиперплоскостью



PIERIAN DATA

Метод опорных векторов

- Разделение пространства гиперплоскостью

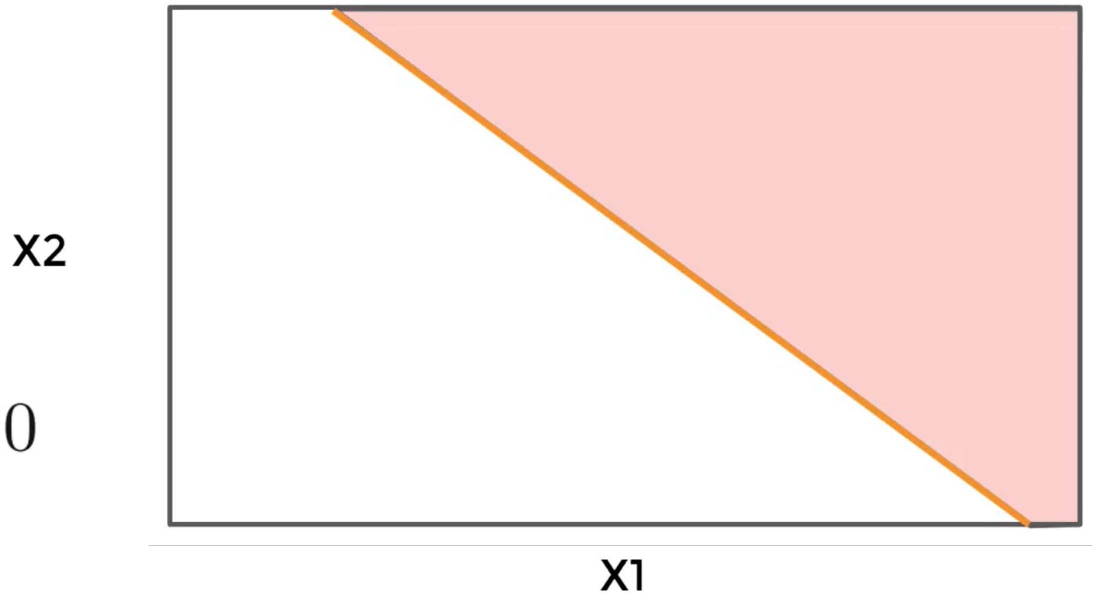


PIERIAN DATA

Метод опорных векторов

- Разделение пространства гиперплоскостью

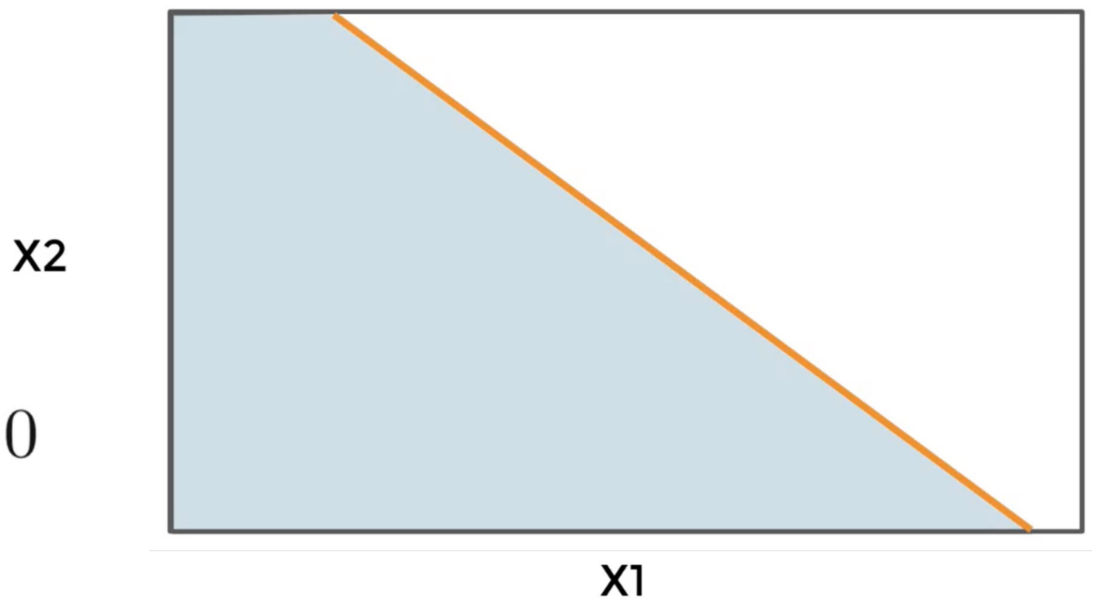
$$\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p > 0$$



Метод опорных векторов

- Разделение пространства гиперплоскостью

$$\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p < 0$$

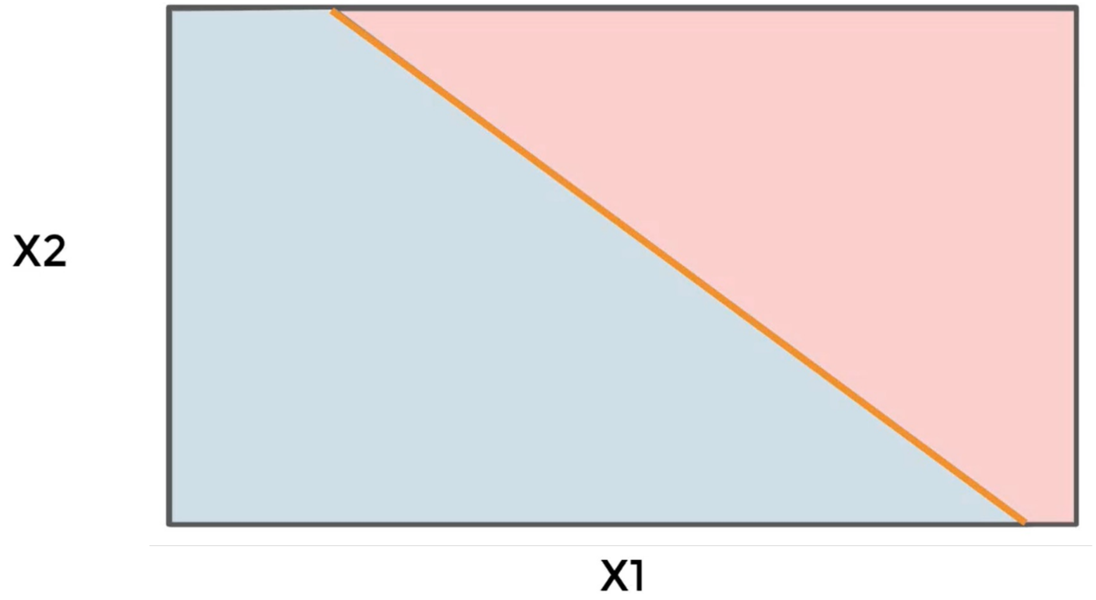


Метод опорных векторов

- Разделение пространства гиперплоскостью

$$\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p > 0$$

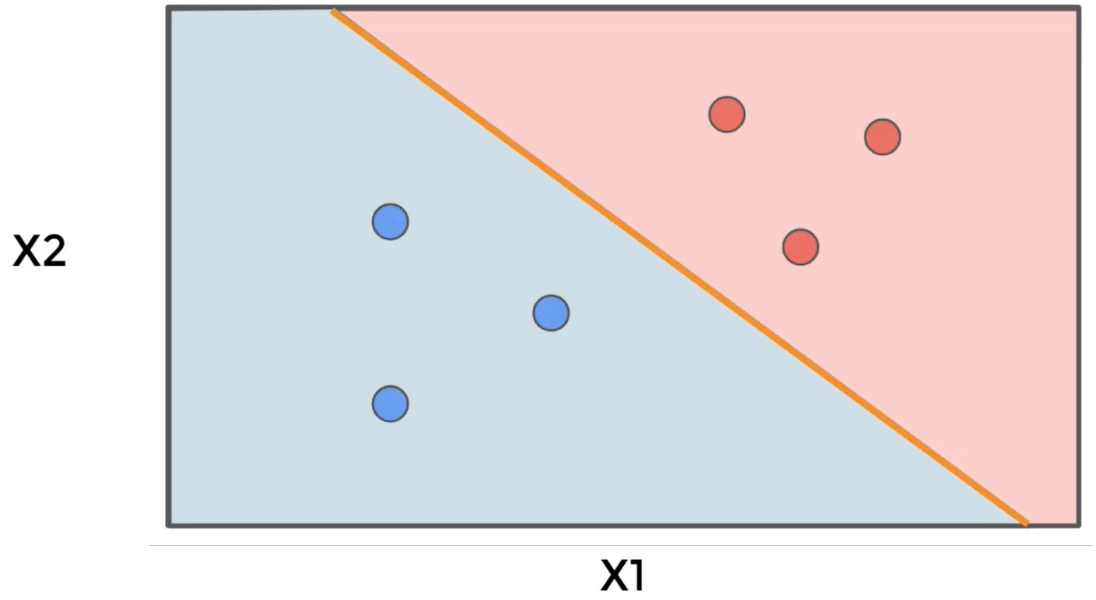
$$\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p < 0$$



Метод опорных векторов

- Точки с данными

$$x_1 = \begin{pmatrix} x_{11} \\ \vdots \\ x_{1p} \end{pmatrix}, \ldots, x_n = \begin{pmatrix} x_{n1} \\ \vdots \\ x_{np} \end{pmatrix}$$

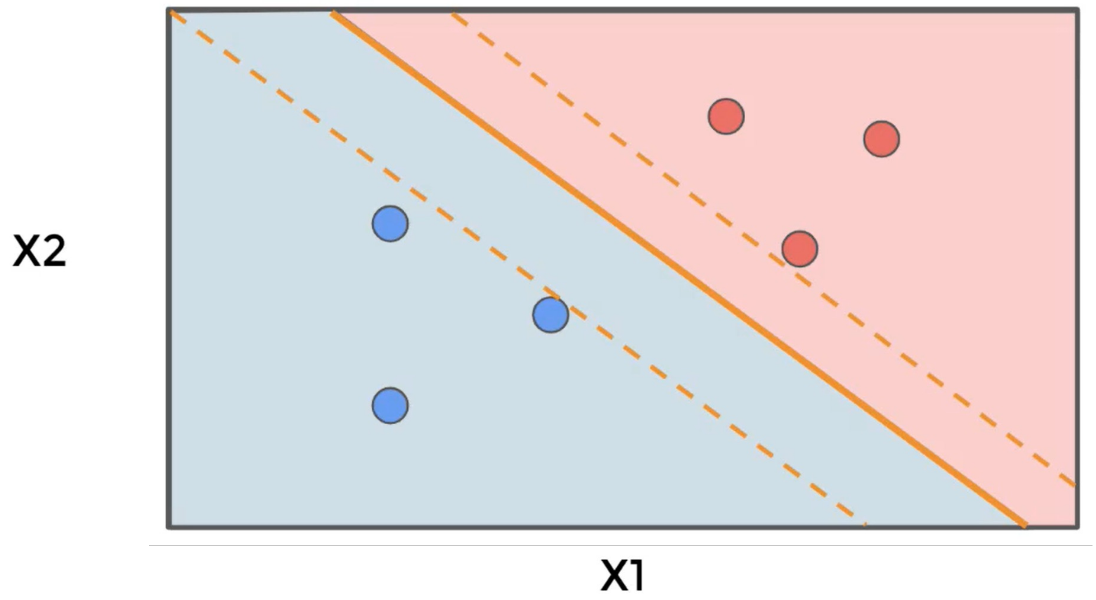


Метод опорных векторов

- Классификатор максимального зазора

Max Margin Classifier

$$\underset{\beta_0,\beta_1,\ldots,\beta_p,M}{\text{максимизируем}} \quad M$$



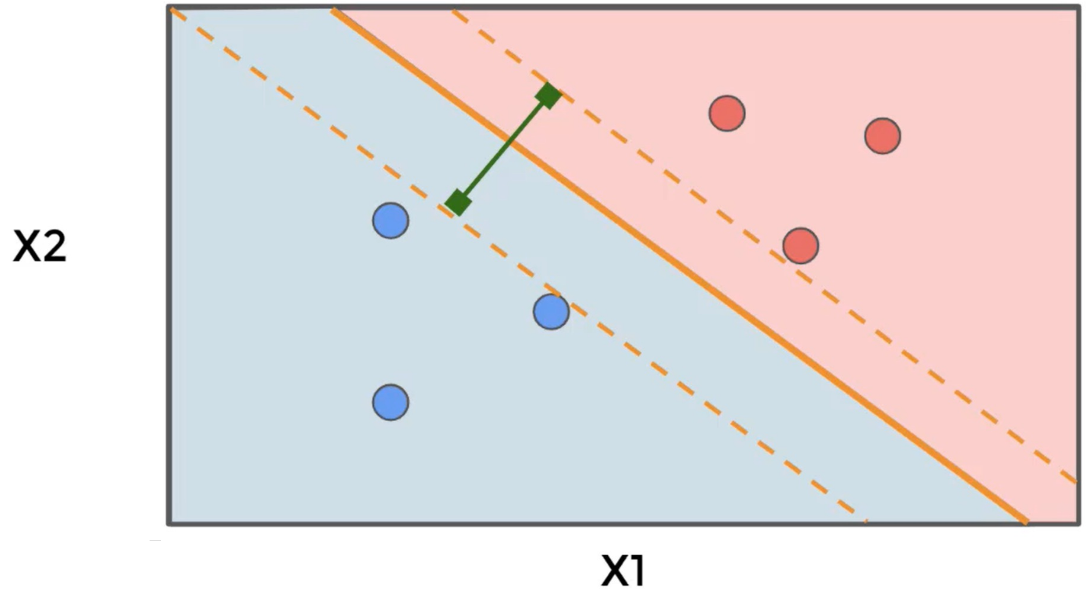
PIERIAN DATA

Метод опорных векторов

- Классификатор максимального зазора

Max Margin Classifier

$$\underset{\beta_0,\beta_1,\ldots,\beta_p,M}{\text{максимизируем}} \quad \boxed{M}$$

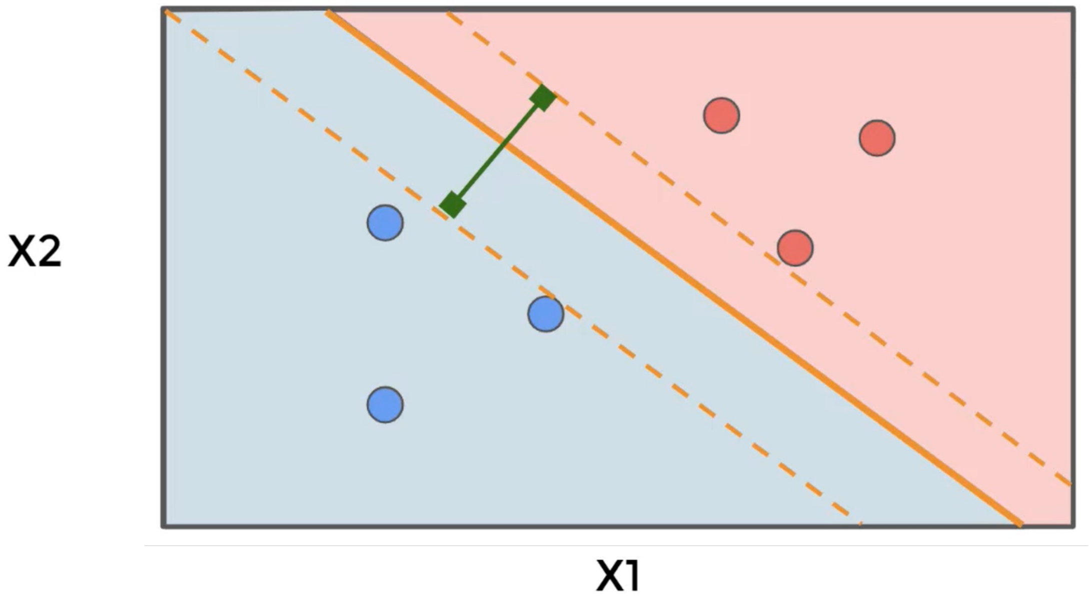


PIERIAN DATA

Метод опорных векторов

Max Margin Classifier

- Классификатор максимального зазора



$$\underset{\beta_0,\beta_1,\ldots,\beta_p,M}{\text{максимизируем}} \quad \boxed{M}$$

$$\text{при условии} \quad \sum_{j=1}^{p} \beta_j^2 = 1$$

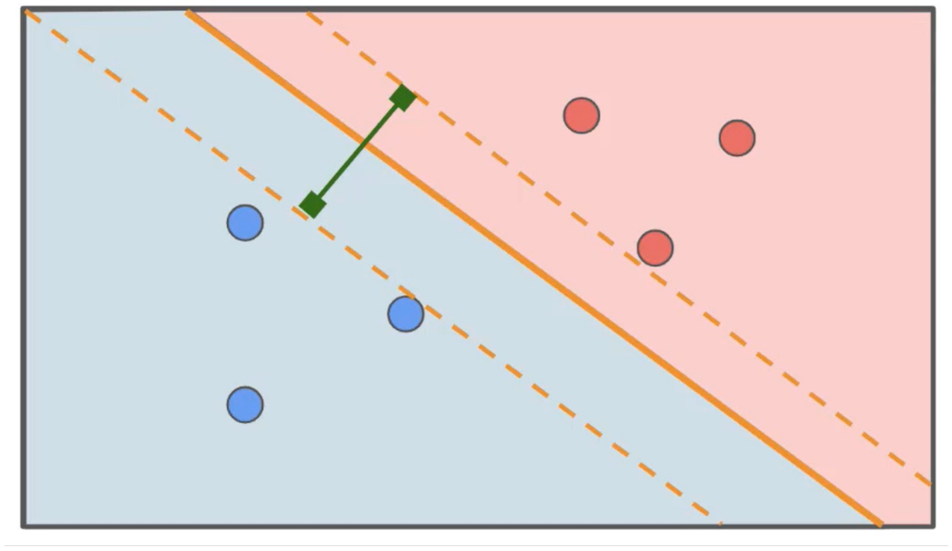
$$y_i(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_p x_{ip}) \geq M \;\; \forall\, i = 1, \ldots, n.$$

Метод опорных векторов

- Классификатор максимального зазора

Max Margin Classifier

$$x_1 = \begin{pmatrix} x_{11} \\ \vdots \\ x_{1p} \end{pmatrix}, \ldots, x_n = \begin{pmatrix} x_{n1} \\ \vdots \\ x_{np} \end{pmatrix}$$



$$\text{максимизируем}_{\beta_0, \beta_1, \ldots, \beta_p, M} \ \boxed{M}$$

$$\text{при условии} \quad \sum_{j=1}^{p} \beta_j^2 = 1$$

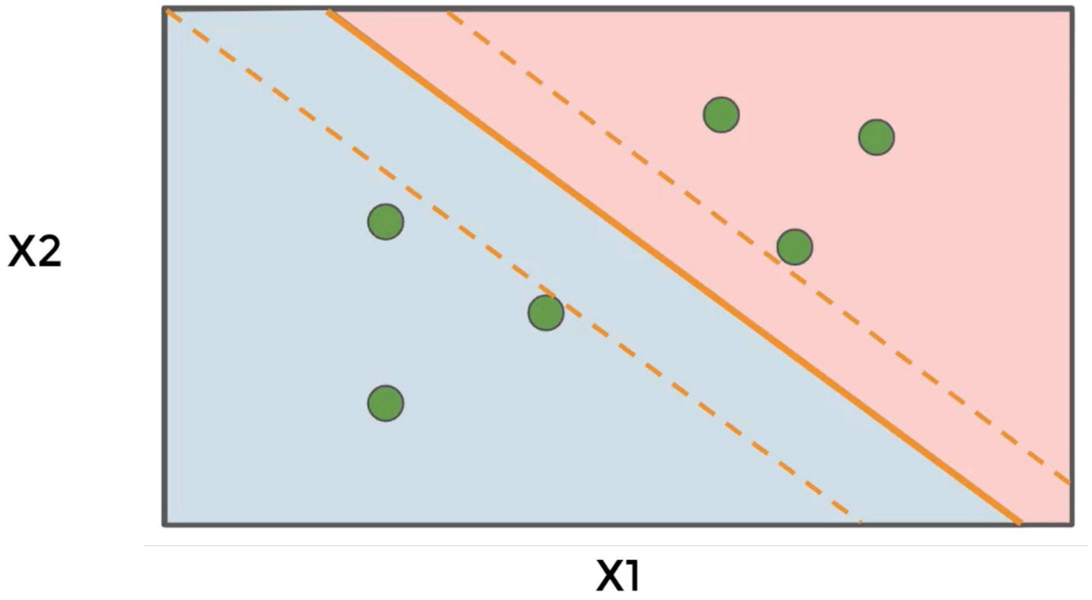
$$y_i(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_p x_{ip}) \geq M \ \ \forall\, i = 1, \ldots, n.$$

Метод опорных векторов

- Классификатор максимального зазора

Max Margin Classifier

$$x_1 = \begin{pmatrix} x_{11} \\ \vdots \\ x_{1p} \end{pmatrix}, \ldots, x_n = \begin{pmatrix} x_{n1} \\ \vdots \\ x_{np} \end{pmatrix}$$



$$\underset{\beta_0,\beta_1,\ldots,\beta_p,M}{\text{максимизируем}} \quad M$$

$$\text{при условии} \quad \sum_{j=1}^{p} \beta_j^2 = 1$$

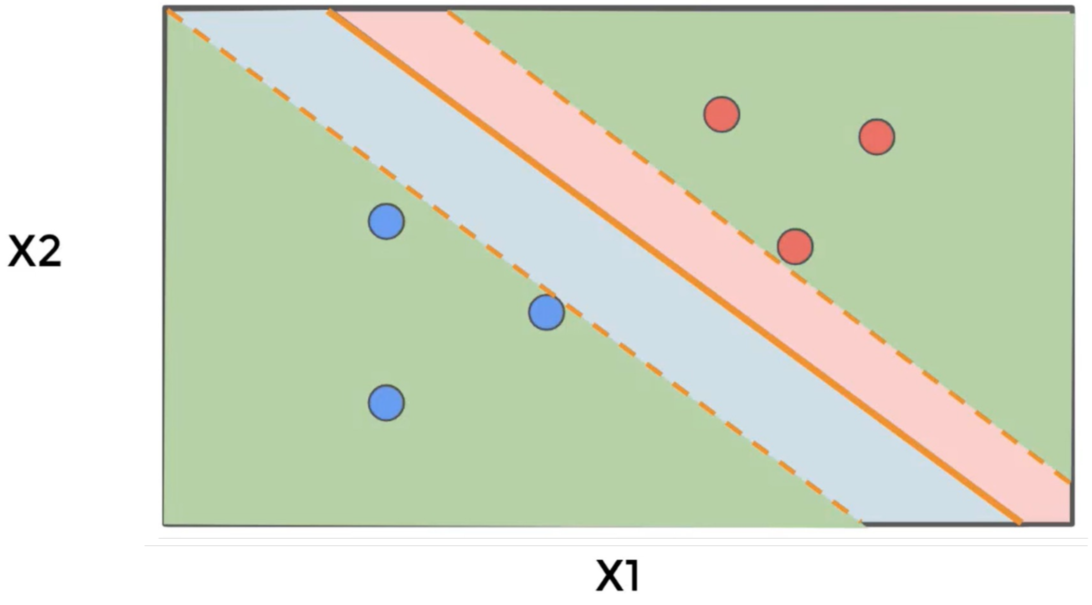
$$y_i(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_p x_{ip}) \geq M \quad \forall\, i = 1, \ldots, n.$$

Метод опорных векторов

Max Margin Classifier

- Классификатор максимального зазора

$$x_1 = \begin{pmatrix} x_{11} \\ \vdots \\ x_{1p} \end{pmatrix}, \ldots, x_n = \begin{pmatrix} x_{n1} \\ \vdots \\ x_{np} \end{pmatrix}$$



максимизируем $\underset{\beta_0, \beta_1, \ldots, \beta_p, M}{M}$

при условии $\sum_{j=1}^{p} \beta_j^2 = 1$

$$\boxed{y_i(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_p x_{ip}) \geq M} \; \forall\, i = 1, \ldots, n.$$

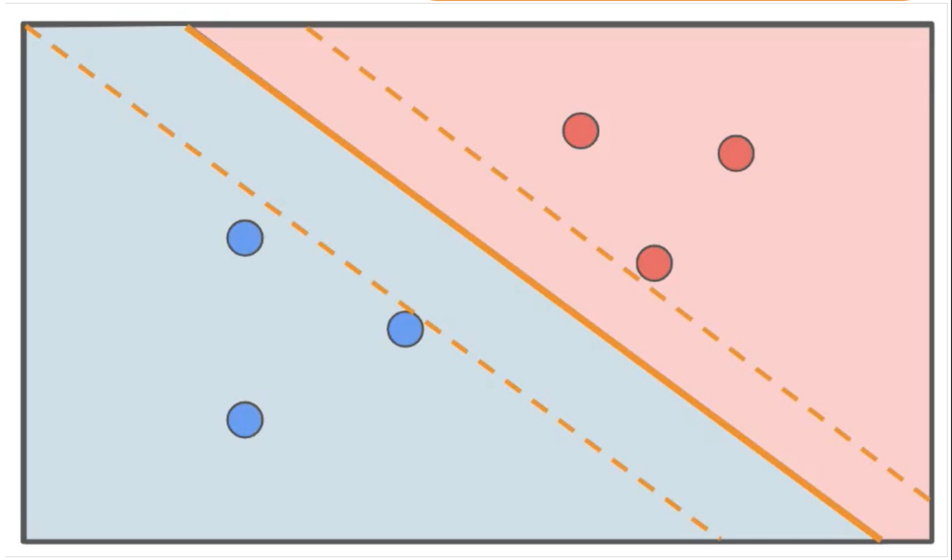
PIERIAN DATA

Метод опорных векторов

Max Margin Classifier

- Классификатор максимального зазора

$$x_1 = \begin{pmatrix} x_{11} \\ \vdots \\ x_{1p} \end{pmatrix}, \ldots, x_n = \begin{pmatrix} x_{n1} \\ \vdots \\ x_{np} \end{pmatrix}$$



максимизируем $\underset{\beta_0,\beta_1,\ldots,\beta_p,M}{M}$

при условии $\boxed{\sum_{j=1}^{p} \beta_j^2 = 1}$

$$y_i(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_p x_{ip}) \geq M \ \ \forall\, i = 1, \ldots, n.$$

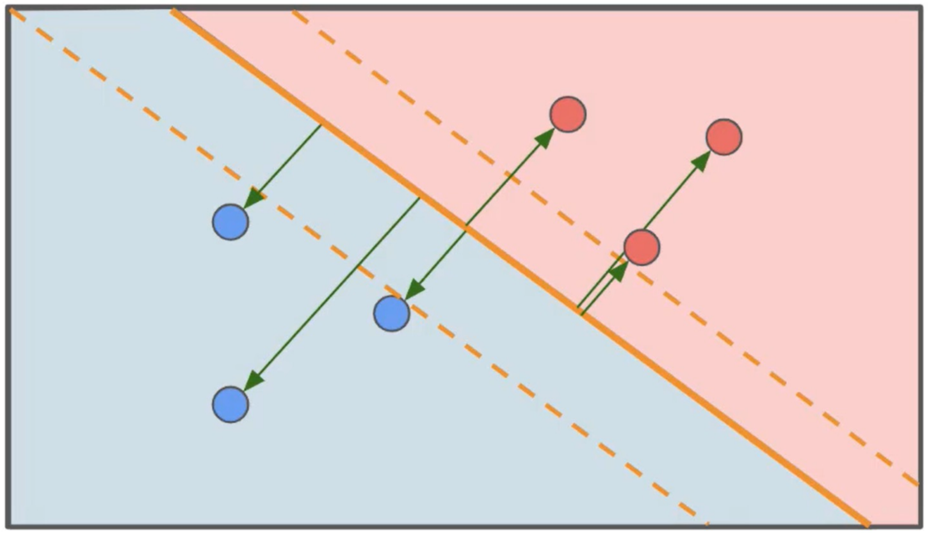
PIERIAN DATA

Метод опорных векторов

- Классификатор максимального зазора

Max Margin Classifier

X2



X1

при условии $\boxed{\sum_{j=1}^{p} \beta_j^2 = 1}$

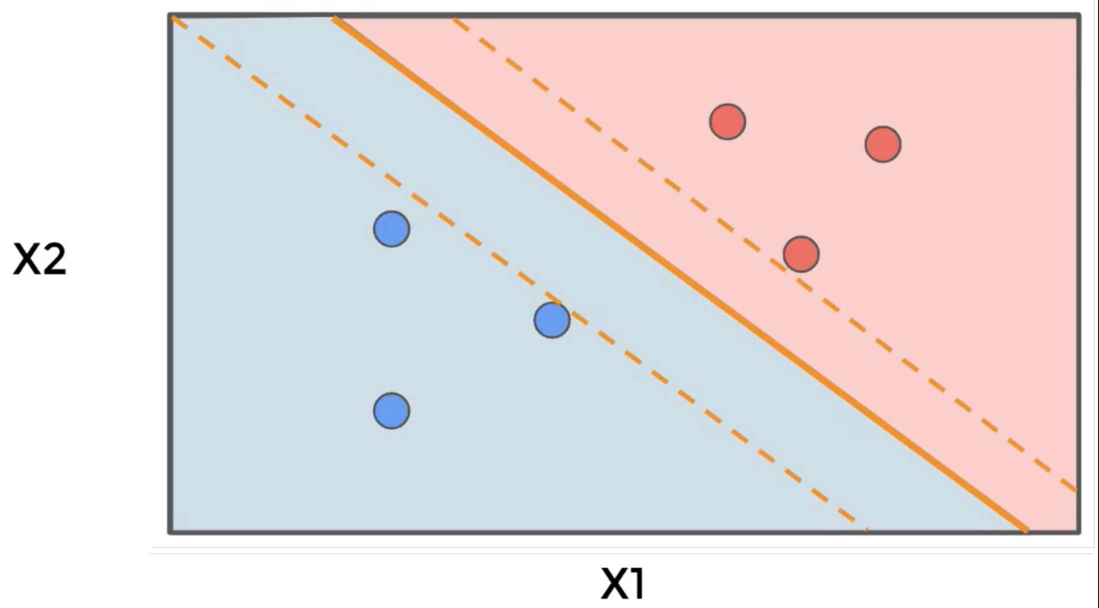
$$y_i(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_p x_{ip}) \geq M \;\; \forall\, i = 1, \ldots, n.$$

Метод опорных векторов

Max Margin Classifier

- Классификатор максимального зазора

$$x_1 = \begin{pmatrix} x_{11} \\ \vdots \\ x_{1p} \end{pmatrix}, \ldots, x_n = \begin{pmatrix} x_{n1} \\ \vdots \\ x_{np} \end{pmatrix}$$



максимизируем $\underset{\beta_0,\beta_1,\ldots,\beta_p,M}{M}$

при условии $\sum_{j=1}^{p} \beta_j^2 = 1$

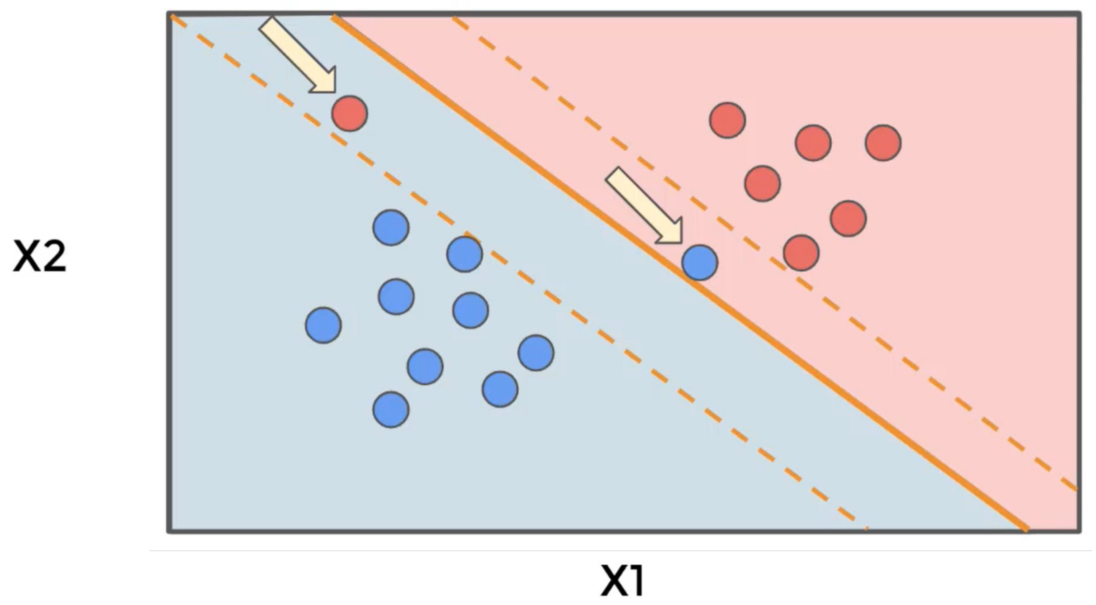
$$y_i(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_p x_{ip}) \geq M \;\; \forall\, i = 1, \ldots, n.$$

Метод опорных векторов

Support Vector Classifier

- Классификатор опорных векторов

$$x_1 = \begin{pmatrix} x_{11} \\ \vdots \\ x_{1p} \end{pmatrix}, \ldots, x_n = \begin{pmatrix} x_{n1} \\ \vdots \\ x_{np} \end{pmatrix}$$



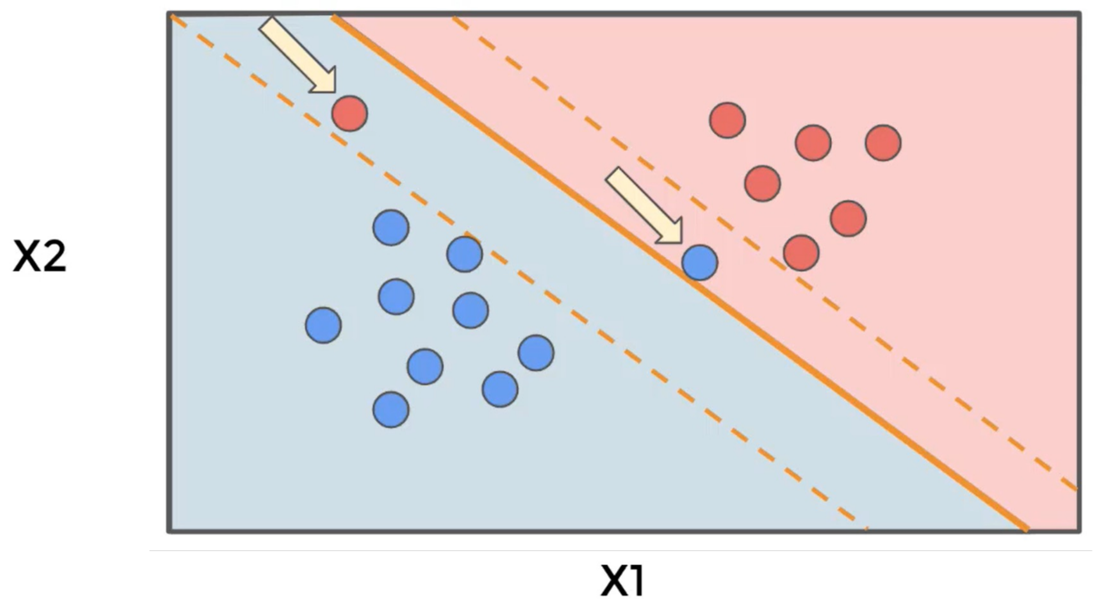
Метод опорных векторов

Support Vector Classifier

- Классификатор опорных векторов

$$\underset{\beta_0,\beta_1,\ldots,\beta_p,\epsilon_1,\ldots,\epsilon_n,M}{\text{максимизируем}} \quad M$$

$$\text{при условии} \quad \sum_{j=1}^{p} \beta_j^2 = 1$$



Метод опорных векторов

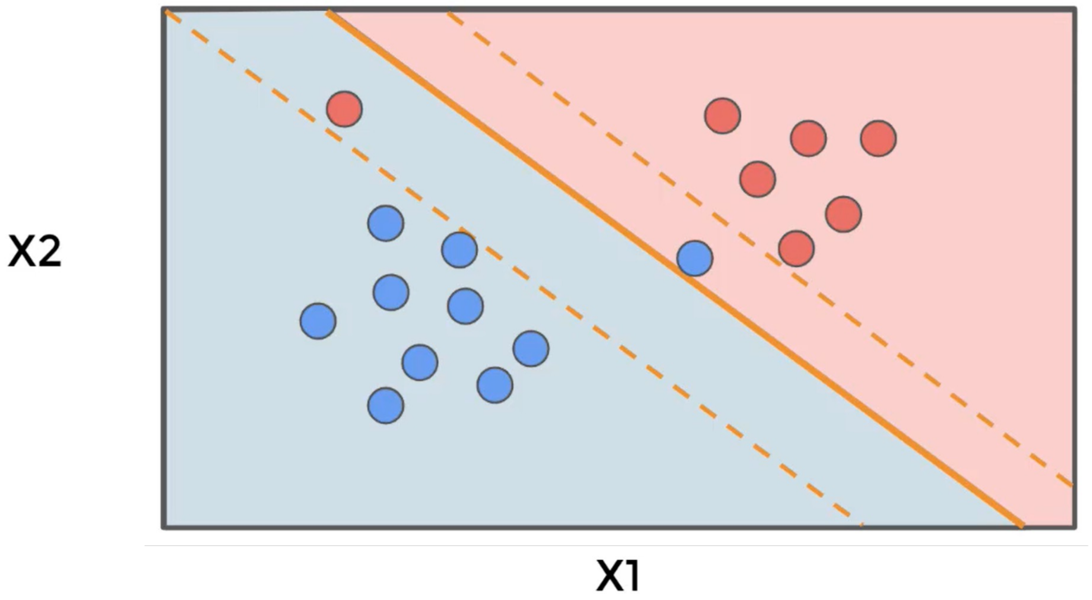
- Классификатор опорных векторов

Support Vector Classifier

$$\underset{\beta_0,\beta_1,\ldots,\beta_p,\epsilon_1,\ldots,\epsilon_n,M}{\text{максимизируем}} \quad M$$

$$\text{при условии} \quad \sum_{j=1}^{p} \beta_j^2 = 1$$

$$\epsilon_i \geq 0, \quad \sum_{i=1}^{n} \epsilon_i \leq C$$



$$y_i(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_p x_{ip}) \geq M(1 - \epsilon_i)$$

Метод опорных векторов

- Классификатор опорных векторов

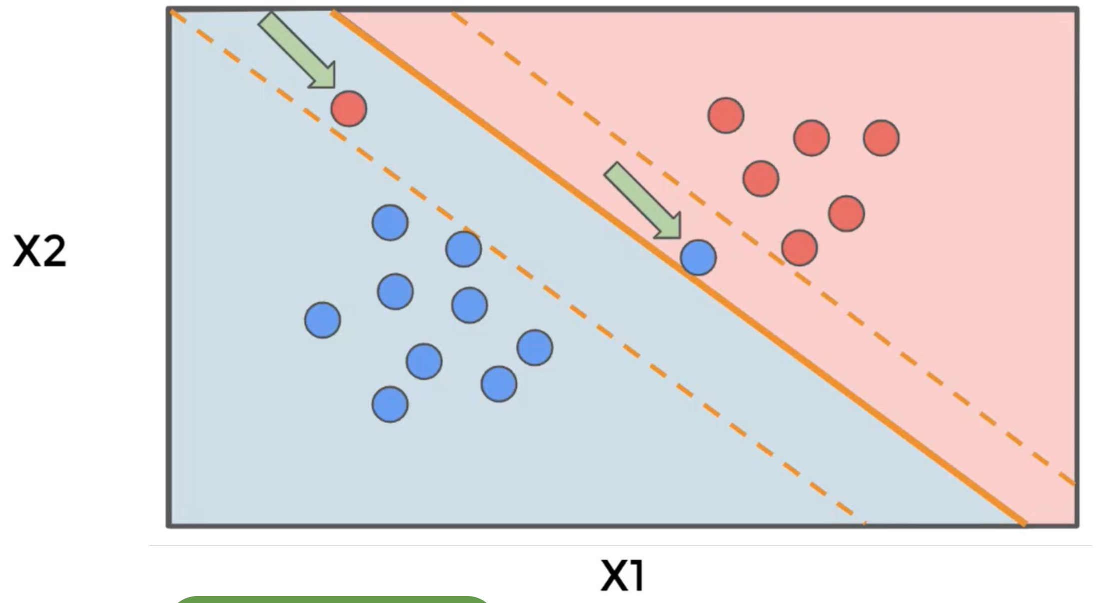
Support Vector Classifier

$$\underset{\beta_0,\beta_1,\ldots,\beta_p,\epsilon_1,\ldots,\epsilon_n,M}{\text{максимизируем}} \quad M$$

$$\text{при условии} \quad \sum_{j=1}^{p} \beta_j^2 = 1$$

$$\boxed{\epsilon_i \geq 0, \quad \sum_{i=1}^{n} \epsilon_i \leq C}$$

$$y_i(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_p x_{ip}) \geq \boxed{M(1 - \epsilon_i)}$$

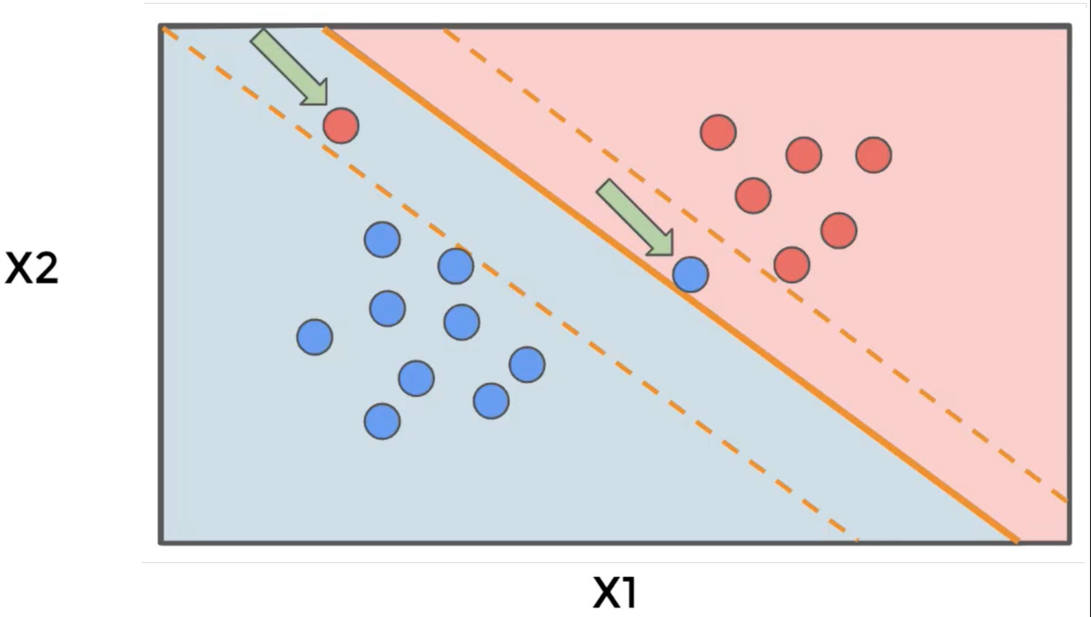


Метод опорных векторов

- Замечание по поводу SVC в Scikit-Learn

C : *float, default=1.0*

Regularization parameter. The strength of the regularization is inversely proportional to C. Must be strictly positive. The penalty is a squared l2 penalty.



$$\epsilon_i \geq 0, \quad \sum_{i=1}^{n} \epsilon_i \leq C$$

$$y_i(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_p x_{ip}) \geq M(1 - \epsilon_i)$$

PIERIAN DATA

Метод опорных векторов

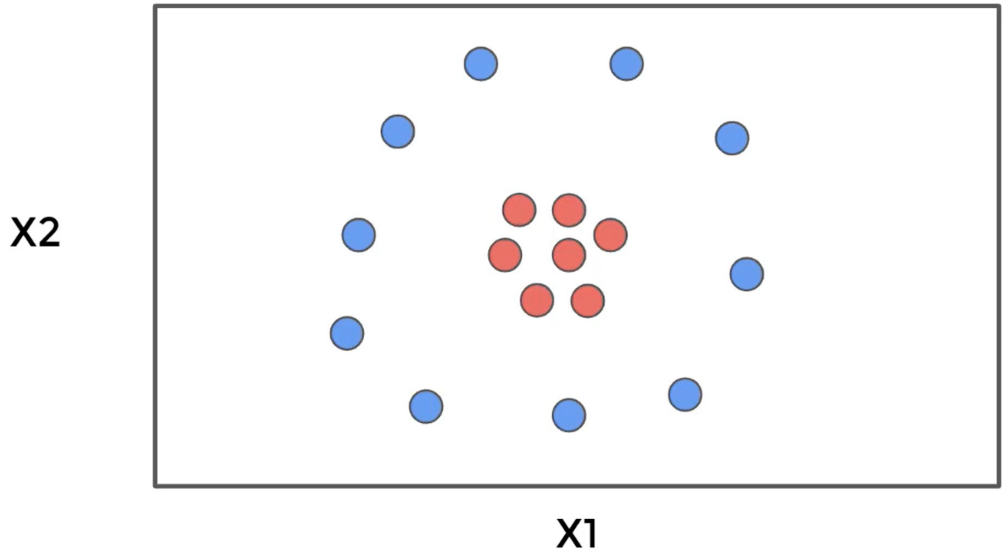
Support Vector Machines

- Метод опорных векторов

$$X_1, X_2, \ldots, X_p,$$

$$\Downarrow$$

$$X_1, X_1^2, X_2, X_2^2, \ldots, X_p, X_p^2$$

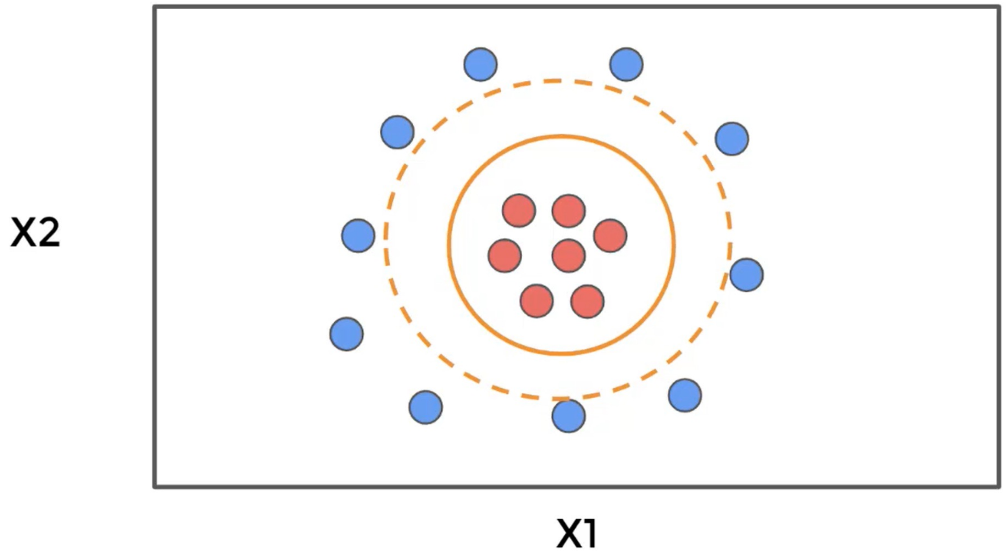


Метод опорных векторов

- Метод опорных векторов

Support Vector Machines

$$\underset{\beta_0,\beta_{11},\beta_{12}\ldots,\beta_{p1},\beta_{p2},\epsilon_1,\ldots,\epsilon_n,M}{\text{максимизируем}} \quad M$$



Метод опорных векторов

Support Vector Machines

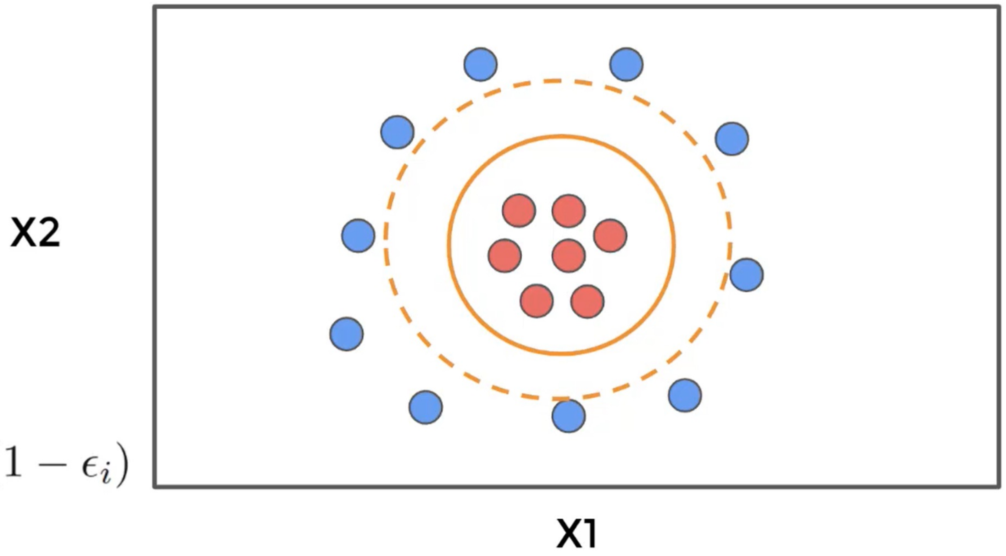
- Метод опорных векторов

$$X_1, X_1^2, X_2, X_2^2, \ldots, X_p, X_p^2$$

максимизируем $\underset{\beta_0,\beta_{11},\beta_{12}\ldots,\beta_{p1},\beta_{p2},\epsilon_1,\ldots,\epsilon_n,M}{M}$

при условии $y_i\left(\beta_0 + \sum_{j=1}^{p}\beta_{j1}x_{ij} + \sum_{j=1}^{p}\beta_{j2}x_{ij}^2\right) \geq M(1-\epsilon_i)$

$$\sum_{i=1}^{n}\epsilon_i \leq C, \quad \epsilon_i \geq 0, \quad \sum_{j=1}^{p}\sum_{k=1}^{2}\beta_{jk}^2 = 1.$$



Метод опорных векторов

Support Vector Machines

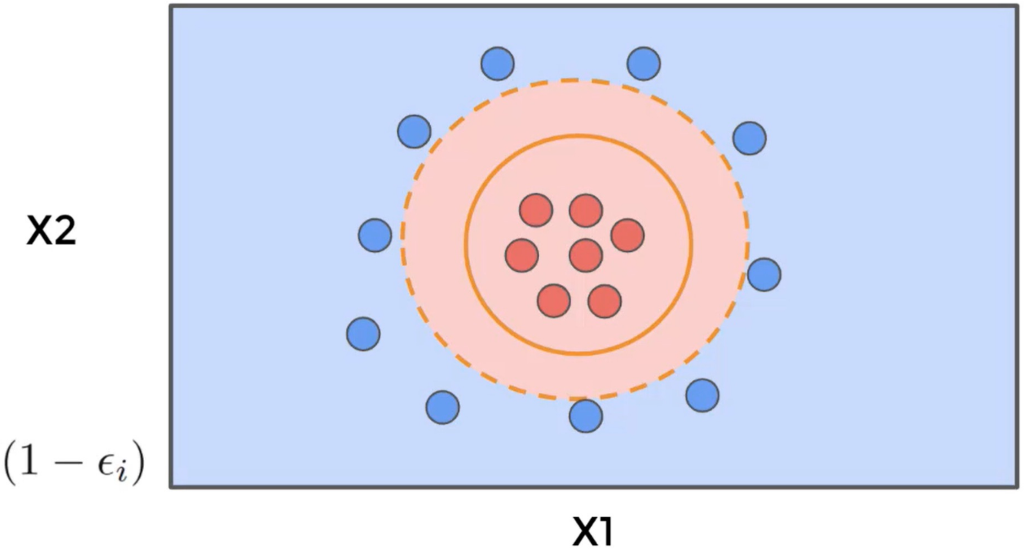
- Метод опорных векторов

$$X_1, X_1^2, X_2, X_2^2, \ldots, X_p, X_p^2$$

максимизируем $\underset{\beta_0,\beta_{11},\beta_{12}\ldots,\beta_{p1},\beta_{p2},\epsilon_1,\ldots,\epsilon_n,M}{} M$

при условии $y_i\left(\beta_0 + \sum_{j=1}^{p}\beta_{j1}x_{ij} + \sum_{j=1}^{p}\beta_{j2}x_{ij}^2\right) \geq M(1-\epsilon_i)$

$$\sum_{i=1}^{n}\epsilon_i \leq C, \quad \epsilon_i \geq 0, \quad \sum_{j=1}^{p}\sum_{k=1}^{2}\beta_{jk}^2 = 1.$$



PIERIAN DATA

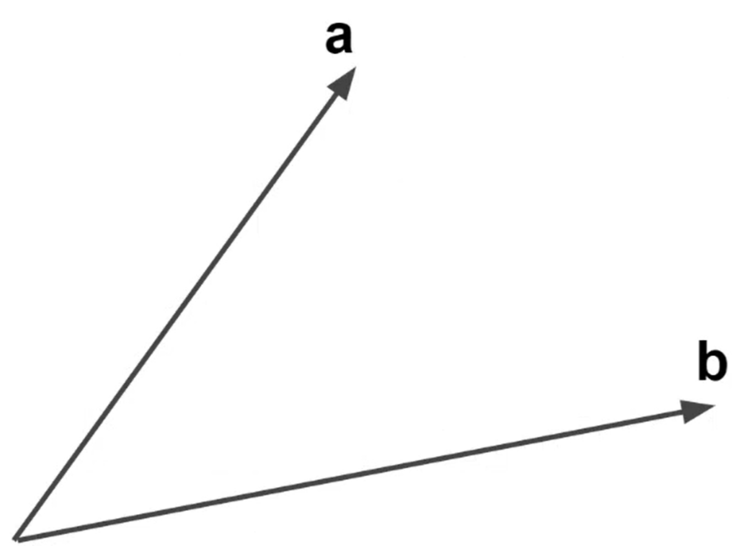
Метод опорных векторов

- Как работать с пространствами очень больших размерностей?
- При росте порядка полиномов растёт и вычислительная трудоёмкость для поиска зазоров.
- Эту сложность помогает решить “kernel trick”, который использует скалярное произведение векторов (dot product)

Метод опорных векторов

- Скалярное произведение векторов (dot product)

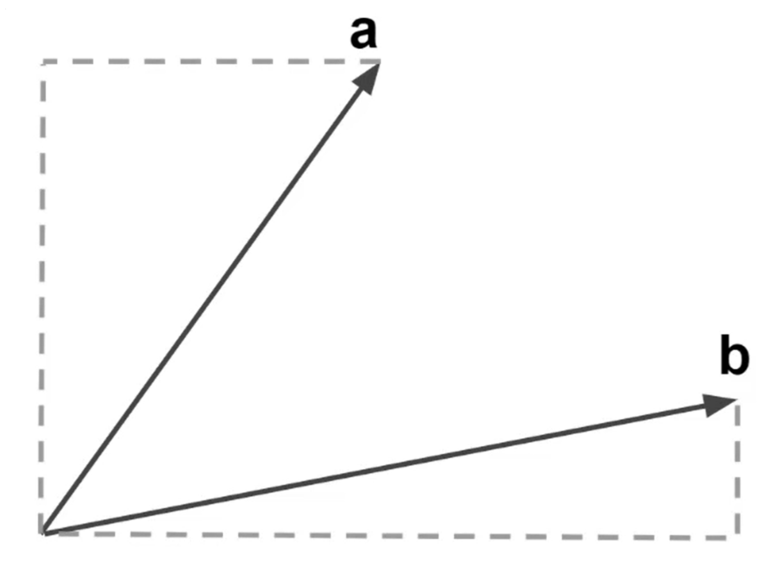
$$\langle a, b \rangle = \sum_{i=1}^{r} a_i b_i$$



Метод опорных векторов

- Скалярное произведение векторов (dot product)

$$\langle a, b \rangle = \sum_{i=1}^{r} a_i b_i$$

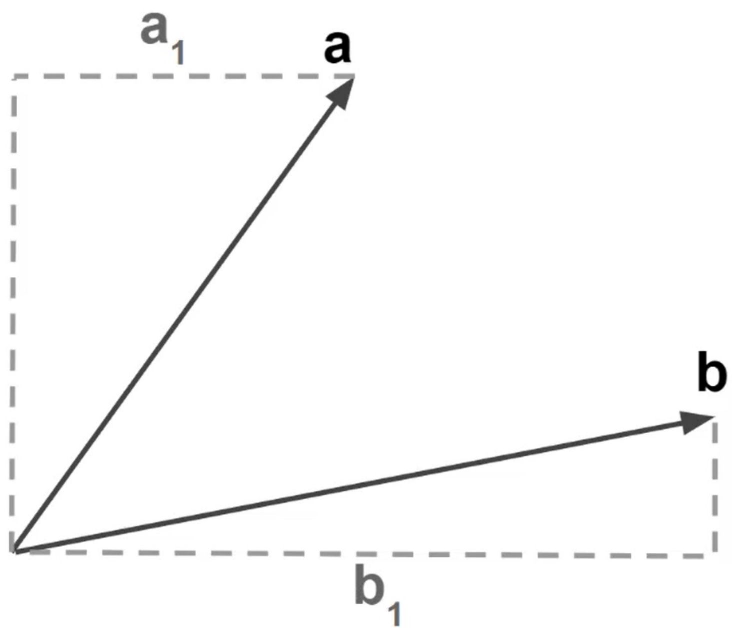


PIERIAN DATA

Метод опорных векторов

- Скалярное произведение векторов (dot product)

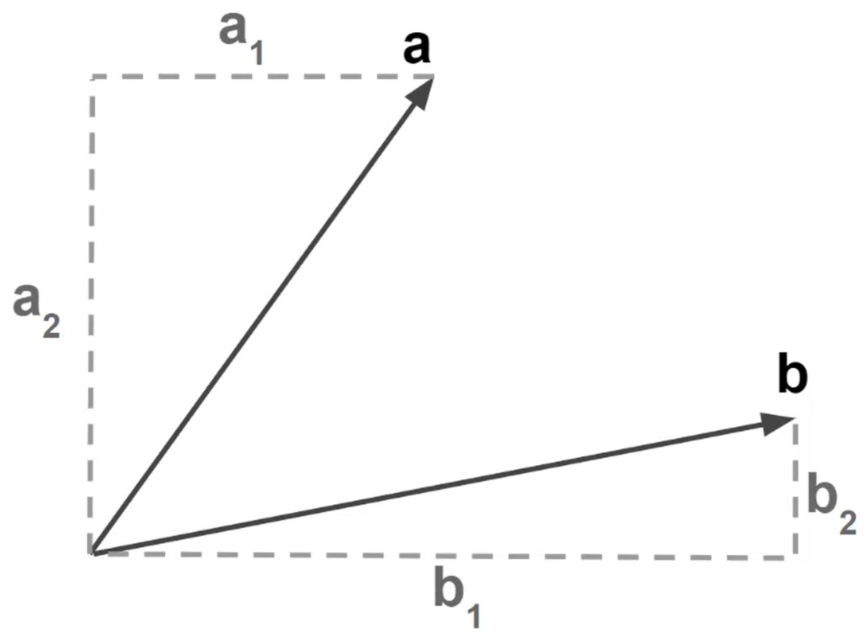
$$\langle a, b\rangle = \sum_{i=1}^{r} a_i b_i$$



Метод опорных векторов

- Скалярное произведение векторов (dot product)

$$\langle a, b \rangle = \sum_{i=1}^{r} a_i b_i$$

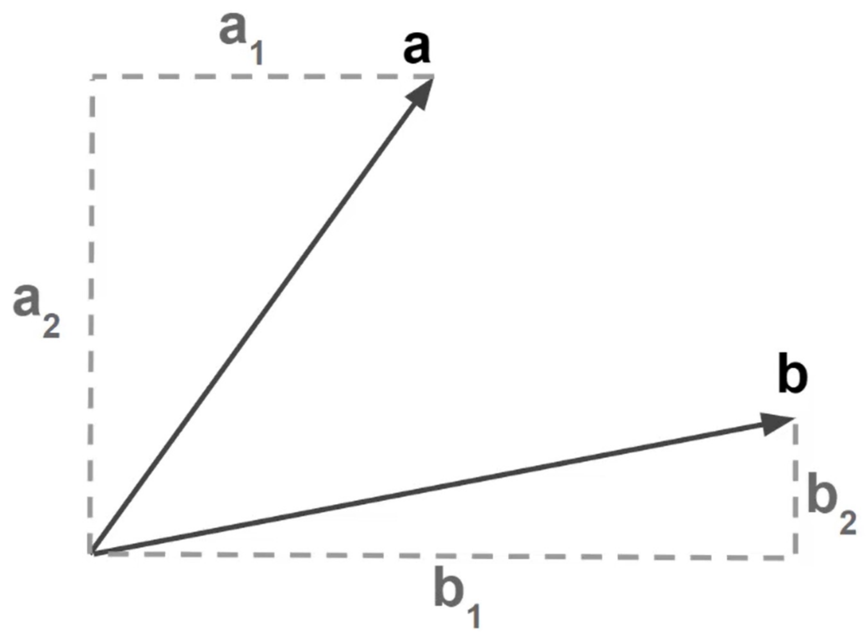


Метод опорных векторов

- Скалярное произведение векторов (dot product)

$$\langle a, b \rangle = \sum_{i=1}^{r} a_i b_i$$

$$a \cdot b = a_1 b_1 + a_2 b_2$$



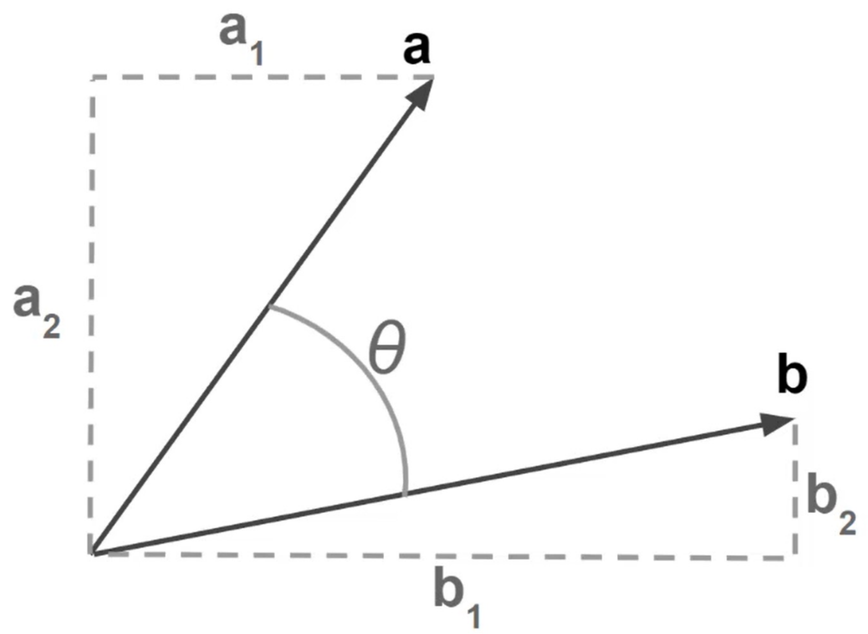
Метод опорных векторов

- Скалярное произведение векторов (dot product)

$$\langle a, b\rangle = \sum_{i=1}^{r} a_i b_i$$

$$a \cdot b = a_1 b_1 + a_2 b_2$$

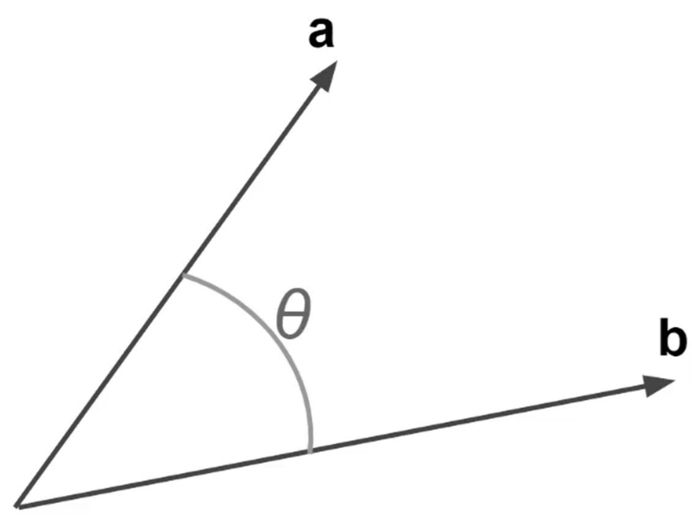
$$a \cdot b = |a||b|cos(\theta)$$



Метод опорных векторов

- Обратите внимание, что скалярное произведение векторов можно представить себе как “похожесть” этих векторов

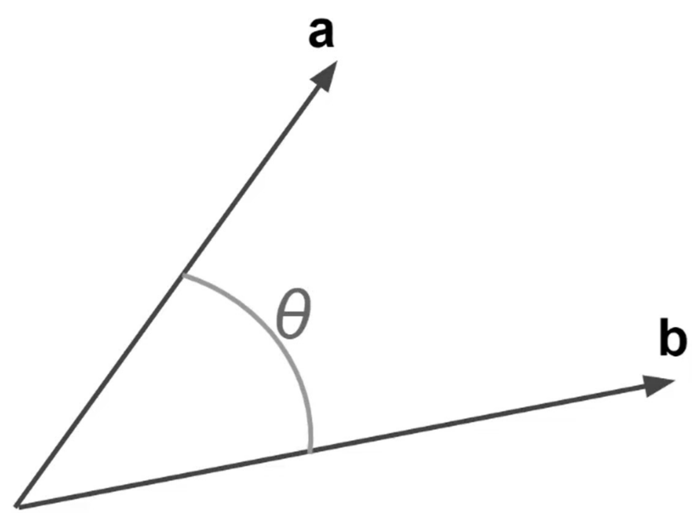
$$a \cdot b = |a||b|cos(\theta)$$



Метод опорных векторов

- $\cos(0°) = 1$
- $\cos(90°) = 0$
- $\cos(180°) = -1$

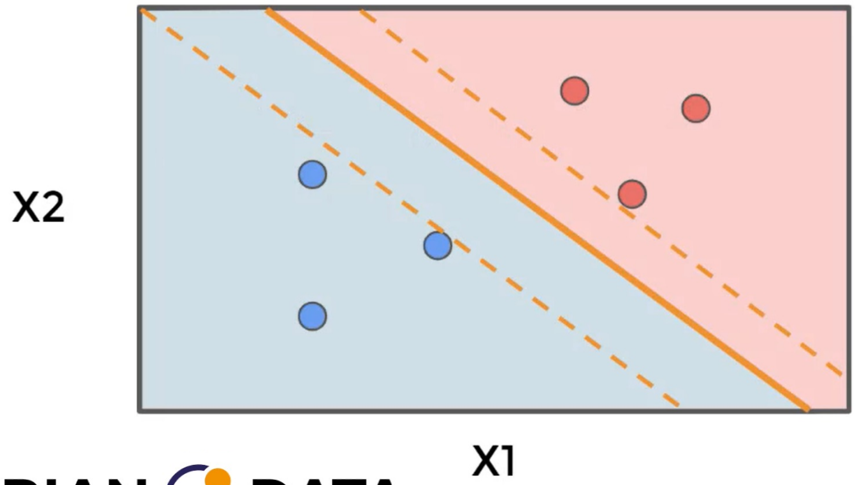
$$a \cdot b = |a||b|cos(\theta)$$



Метод опорных векторов

- Линейный классификатор можно переписать так:

$$f(x) = \beta_0 + \boxed{\sum_{i=1}^{n} \alpha_i \langle x, x_i \rangle}$$



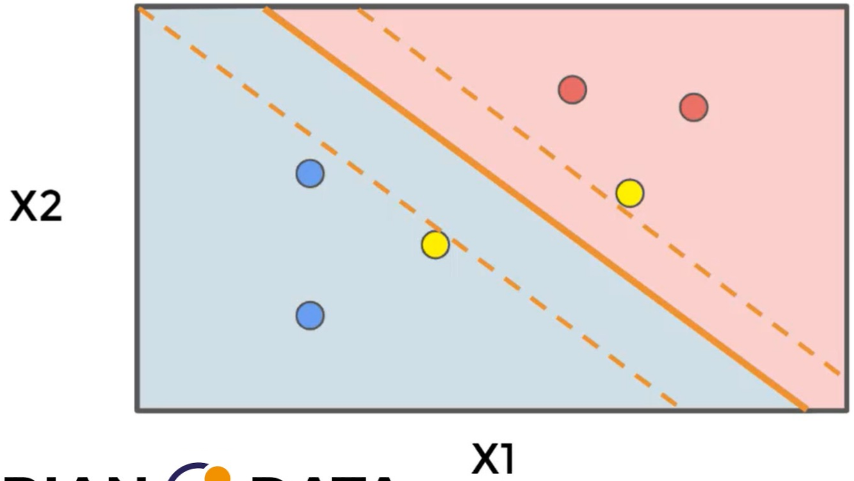
Вычисляем скалярные произведения всех пар обучающих наблюдений

PIERIAN DATA

Метод опорных векторов

- Линейный классификатор можно переписать так:

$$f(x) = \beta_0 + \sum_{i=1}^{n} \boxed{\alpha_i} \langle x, x_i \rangle$$



Только не-нули для опорных векторов

Метод опорных векторов

- Линейный классификатор можно переписать так:

$$f(x) = \beta_0 + \sum_{i=1}^{n} \alpha_i \langle x, x_i \rangle$$

$$f(x) = \beta_0 + \sum_{i \in \mathcal{S}} \alpha_i \langle x, x_i \rangle$$

Метод опорных векторов

- Функция ядра:

$$K(x_i, x_{i'}) = \sum_{j=1}^{p} x_{ij} x_{i'j}$$

Ядро - функция, численно описывающая похожесть двух наблюдений

Метод опорных векторов

- Функция ядра:

$$K(x_i, x_{i'}) = \sum_{j=1}^{p} x_{ij} x_{i'j} \qquad f(x) = \beta_0 + \sum_{i \in \mathcal{S}} \alpha_i \langle x, x_i \rangle$$

Метод опорных векторов

- Функция ядра:

$$K(x_i, x_{i'}) = \boxed{\sum_{j=1}^{p} x_{ij} x_{i'j}} \qquad f(x) = \beta_0 + \sum_{i \in \mathcal{S}} \alpha_i \boxed{\langle x, x_i \rangle}$$

$$\boxed{\langle a, b \rangle = \sum_{i=1}^{r} a_i b_i}$$

Метод опорных векторов

- Полиномиальная функция ядра:

$$\boxed{K(x_i, x_{i'})} = \left(1 + \sum_{j=1}^{p} x_{ij} x_{i'j}\right)^d$$

$$f(x) = \beta_0 + \sum_{i \in \mathcal{S}} \alpha_i \boxed{\langle x, x_i \rangle}$$

$$\boxed{\langle a, b \rangle = \sum_{i=1}^{r} a_i b_i}$$

Метод опорных векторов

- Радиальная базисная функция ядра (radial basis kernel):

$$\boxed{K(x_i, x_{i'})} = \exp(-\gamma \sum_{j=1}^{p}(x_{ij} - x_{i'j})^2)$$

$$f(x) = \beta_0 + \sum_{i \in \mathcal{S}} \alpha_i \boxed{\langle x, x_i \rangle}$$

$$\boxed{\langle a, b \rangle = \sum_{i=1}^{r} a_i b_i}$$

Метод опорных векторов

- Применение функций ядра ещё называют словосочетанием “kernel trick”.
- Ядра позволяют избежать вычислений в увеличенном пространстве признаков, выполняя только вычисления для каждой различной пары обучающих точек (подробнее см. пункт 9.3.2 в ISLR).

Метод опорных векторов

- Ранее мы видели, что скалярное произведение можно представить себе как меру похожести между векторами.
- Функции ядра можно представить себе как меру похожести между исходным пространством признаков и увеличенным пространством признаков.

Метод опорных векторов

- Теперь, когда мы посмотрели на теорию, время перейти к написанию кода!

Метод опорных векторов (support vector machines)

Классификация в Scikit-Learn – часть 1

PIERIAN DATA

Метод опорных векторов

- Мы будем использовать функцию, которая находится в файле .py в папке для метода опорных векторов.

Метод опорных векторов (support vector machines)

Классификация в Scikit-Learn – часть 2

PIERIAN DATA

Метод опорных векторов (support vector machines)

Регрессия в Scikit-Learn

PIERIAN DATA

Метод опорных векторов

Упражнения

PIERIAN DATA

Метод опорных векторов

- Сможем ли мы определить подделку вина, основываясь на результатах химического анализа?

PIERIAN DATA

Метод опорных векторов

- Можно ли вообще решить эту задачу? Возможно ли отличить настоящее вино от поддельного?

Метод опорных векторов

Решения для упражнений

PIERIAN DATA