

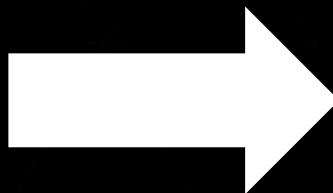
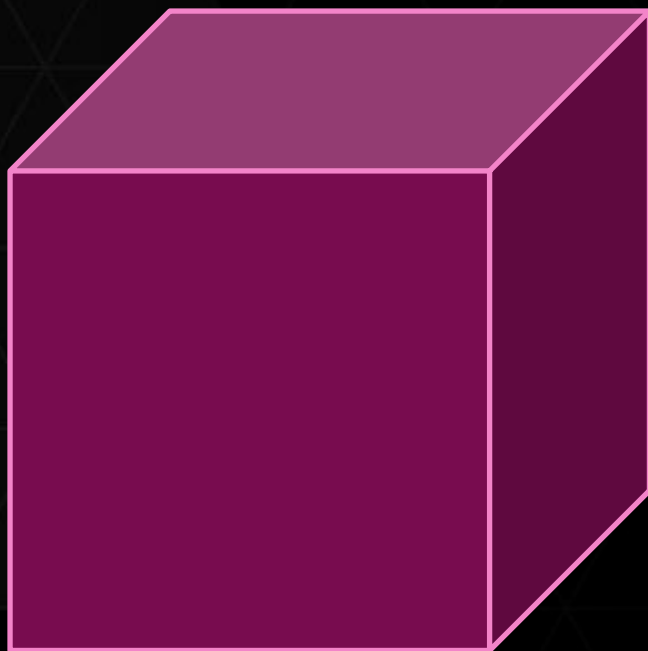


# ОБУЧЕНИЕ ЭМБЕДДИНГОВ

Дмитрий Коробченко

# EMBEDDING (ВЛОЖЕНИЕ)

Объект



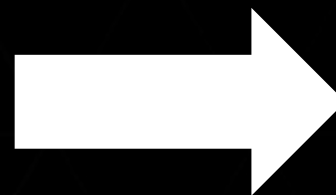
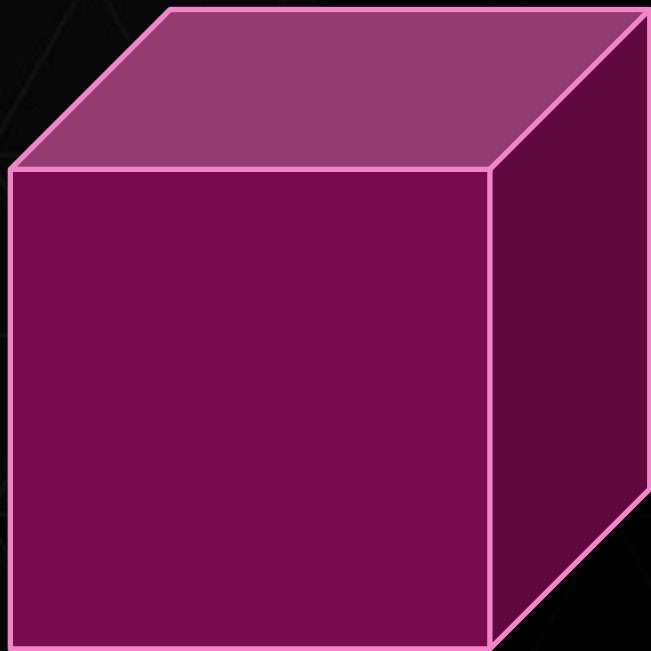
Вектор

$$\begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_N \end{bmatrix}$$

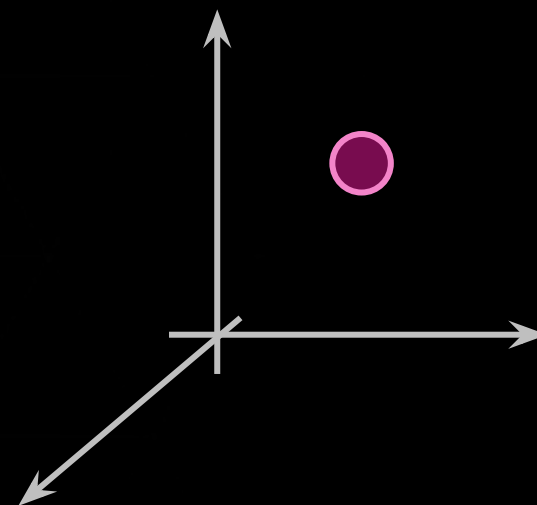
# EMBEDDING (ВЛОЖЕНИЕ)

Вложение в N-мерное

Объект



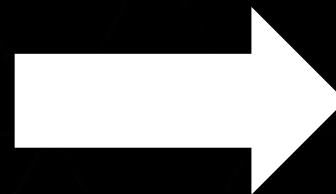
Вектор  
(Эмбе́ддинг)



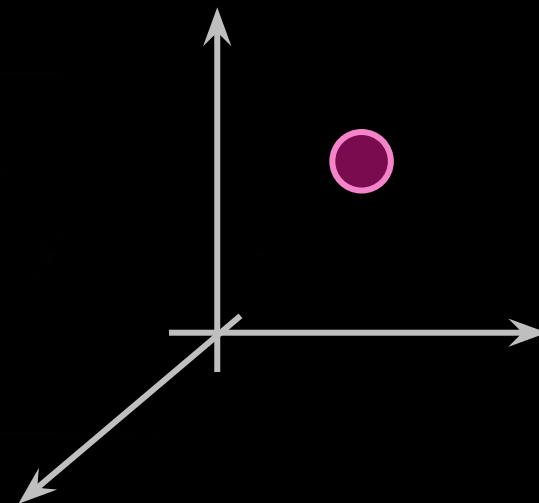


# EMBEDDING (ВЛОЖЕНИЕ)

Изображение

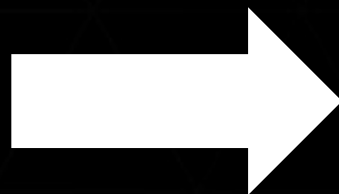
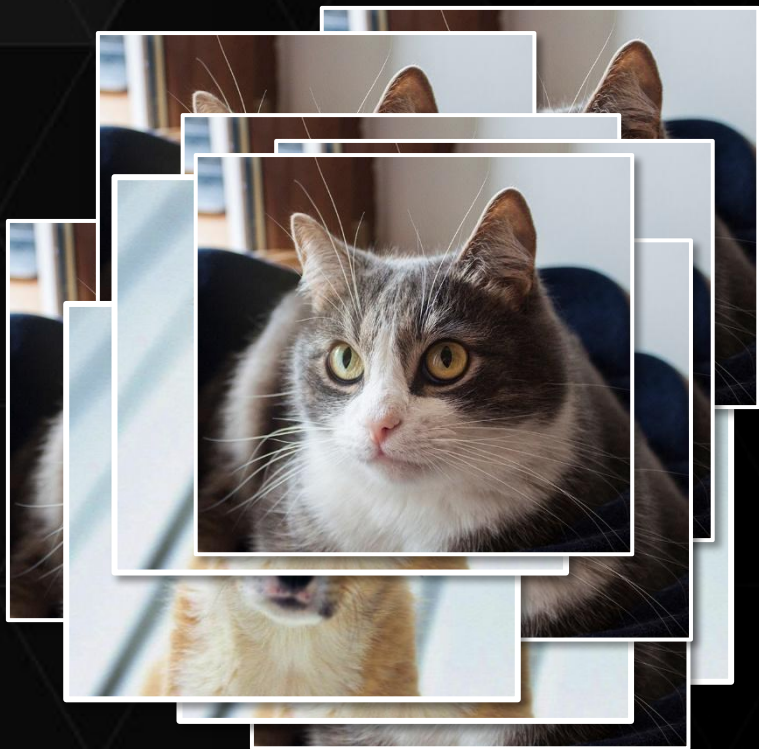


Вектор  
(Эмбе́ддинг)

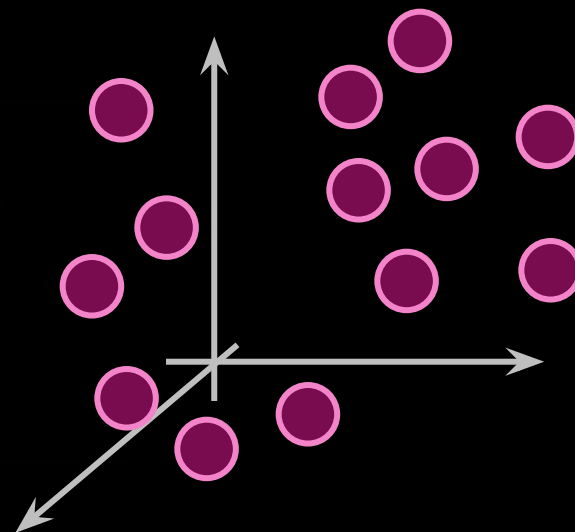


# EMBEDDING (ВЛОЖЕНИЕ)

Изображения



Векторы  
(Эмбединги)

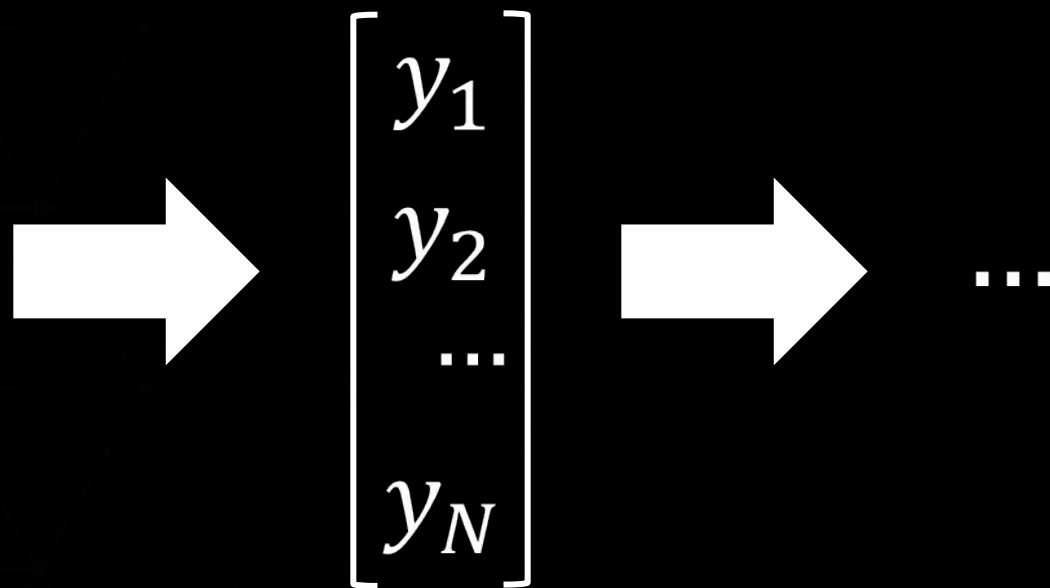


# ЭМБЕДДИНГ КАК ДЕСКРИПТОР

Изображение

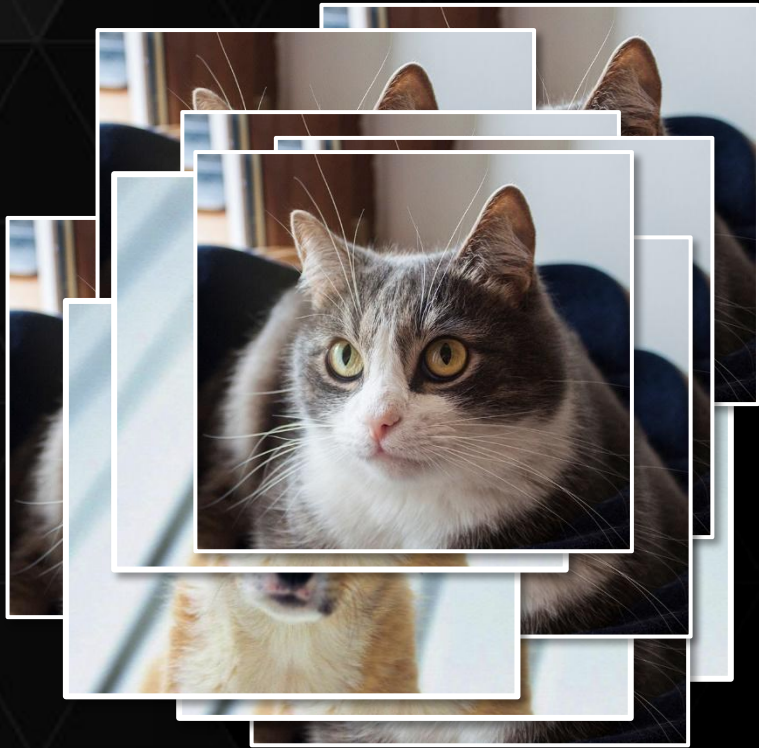


Эмбеддинг



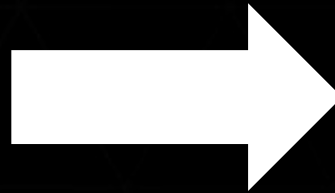
# КЛАСТЕРИЗАЦИЯ ОБЪЕКТОВ

Изображения

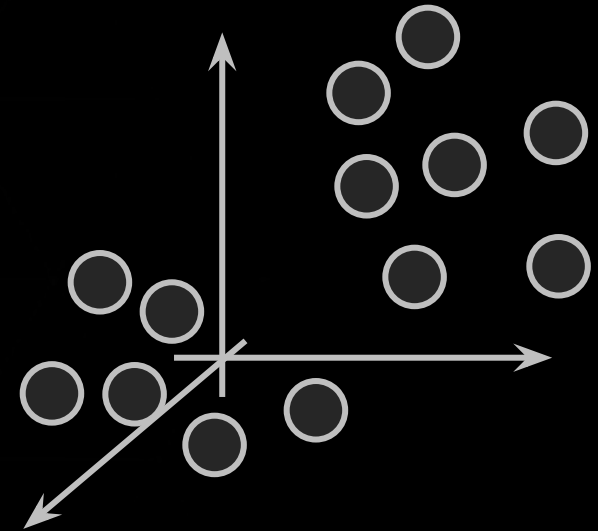


# КЛАСТЕРИЗАЦИЯ ОБЪЕКТОВ

Изображения



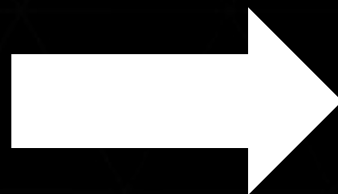
Эмбединги



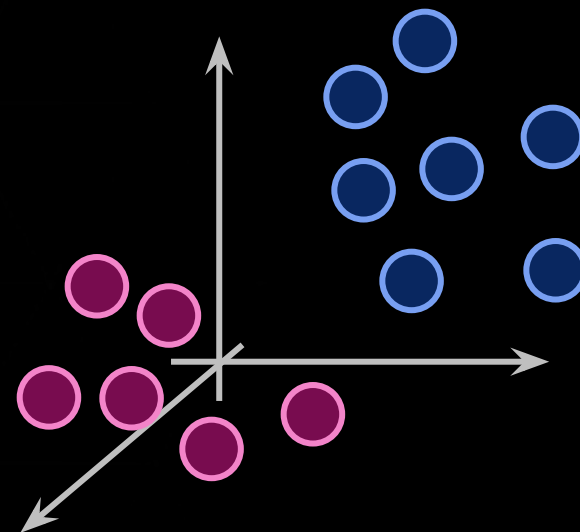


# КЛАСТЕРИЗАЦИЯ ОБЪЕКТОВ

Изображения



Эмбединги  
+ Кластеризация



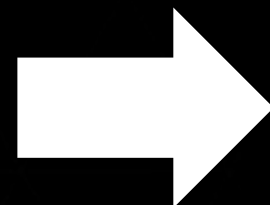
# РАССТОЯНИЕ МЕЖДУ ОБЪЕКТАМИ

Два изображения

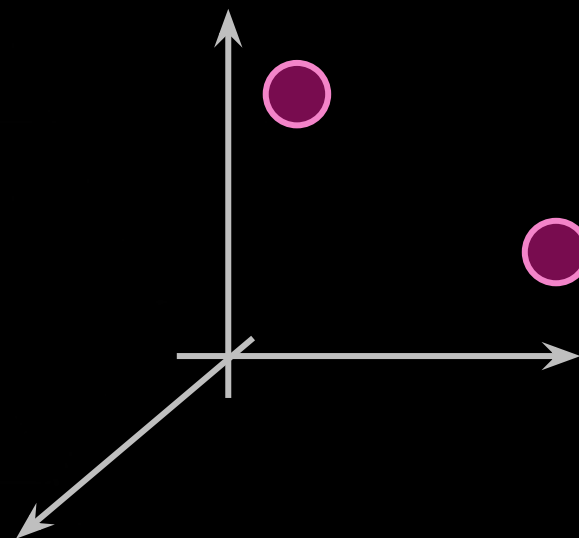


# РАССТОЯНИЕ МЕЖДУ ОБЪЕКТАМИ

Два изображения

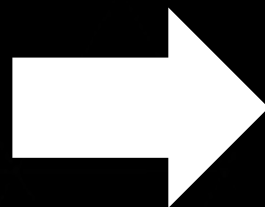


Эмбединги

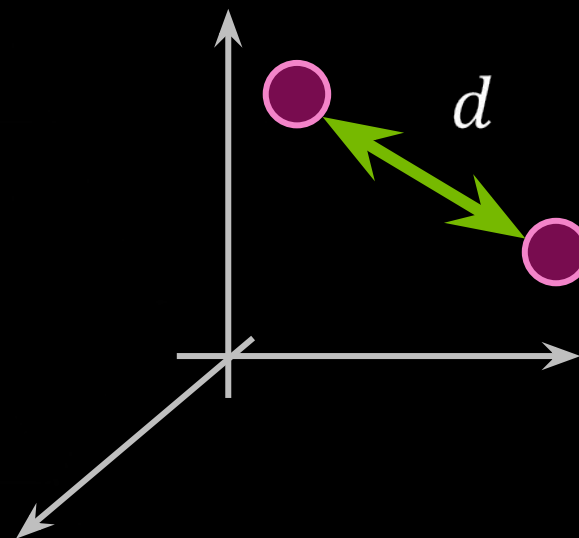


# РАССТОЯНИЕ МЕЖДУ ОБЪЕКТАМИ

Два изображения

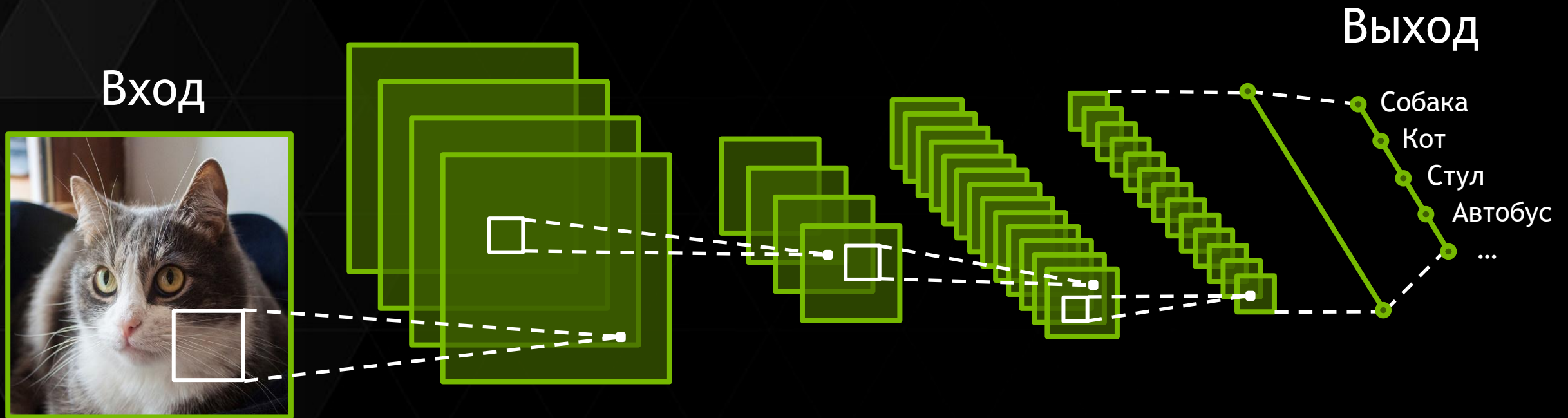


Эмбединги

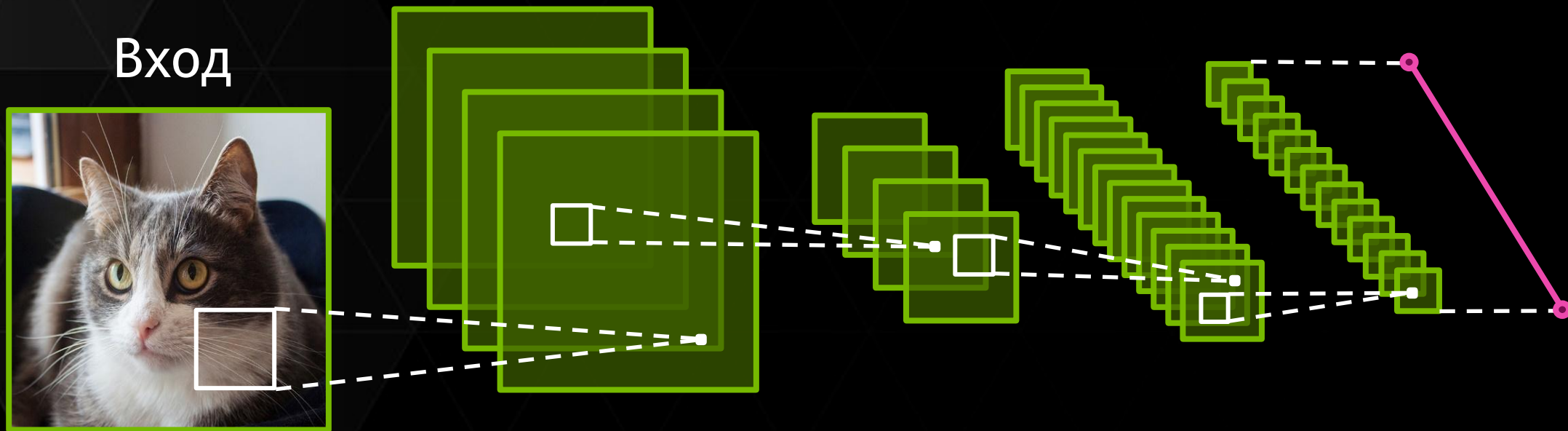




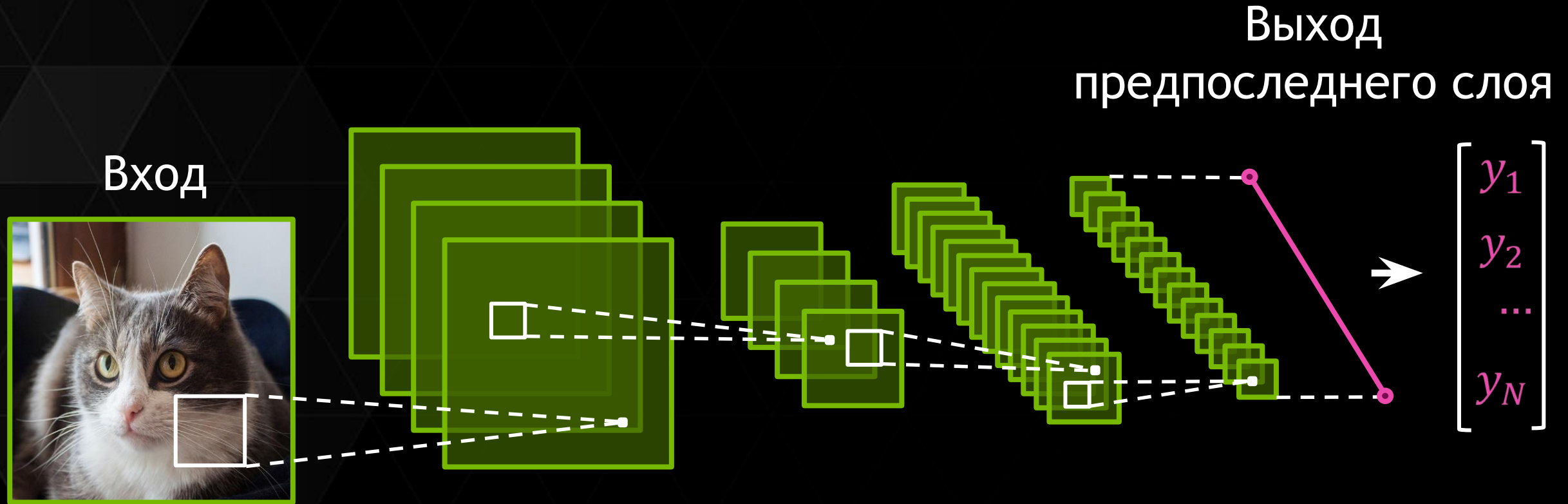
# СВЁРТОЧНАЯ НЕЙРОННАЯ СЕТЬ



# СВЁРТОЧНАЯ НЕЙРОННАЯ СЕТЬ

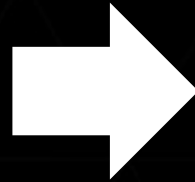


# СВЁРТОЧНАЯ НЕЙРОННАЯ СЕТЬ

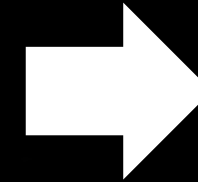
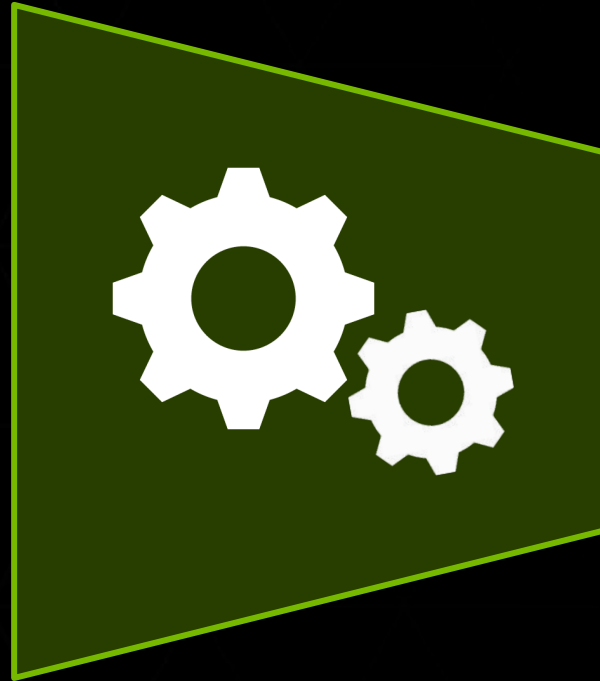


# СВЁРТОЧНАЯ НЕЙРОННАЯ СЕТЬ

Изображение



Произвольная  
CNN

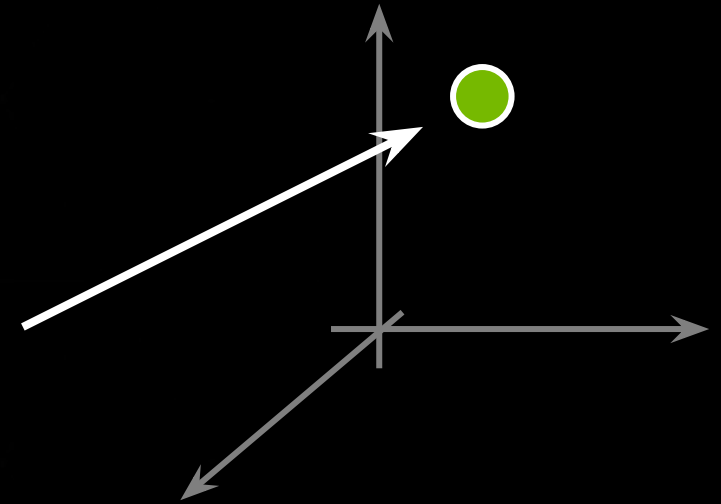
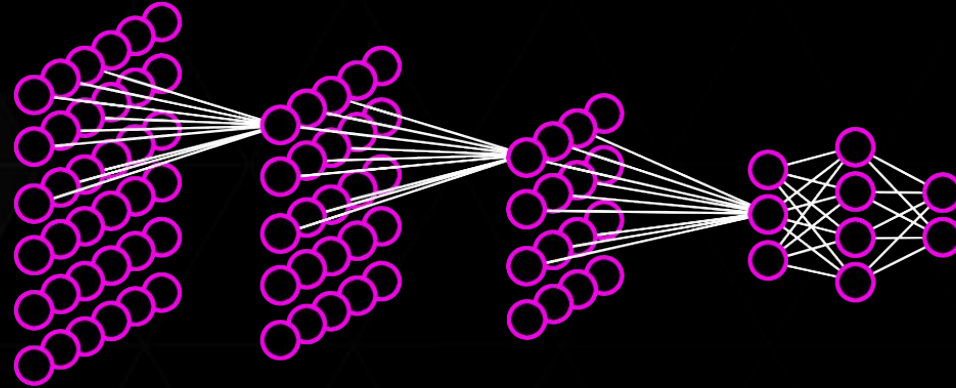


Эмбединг

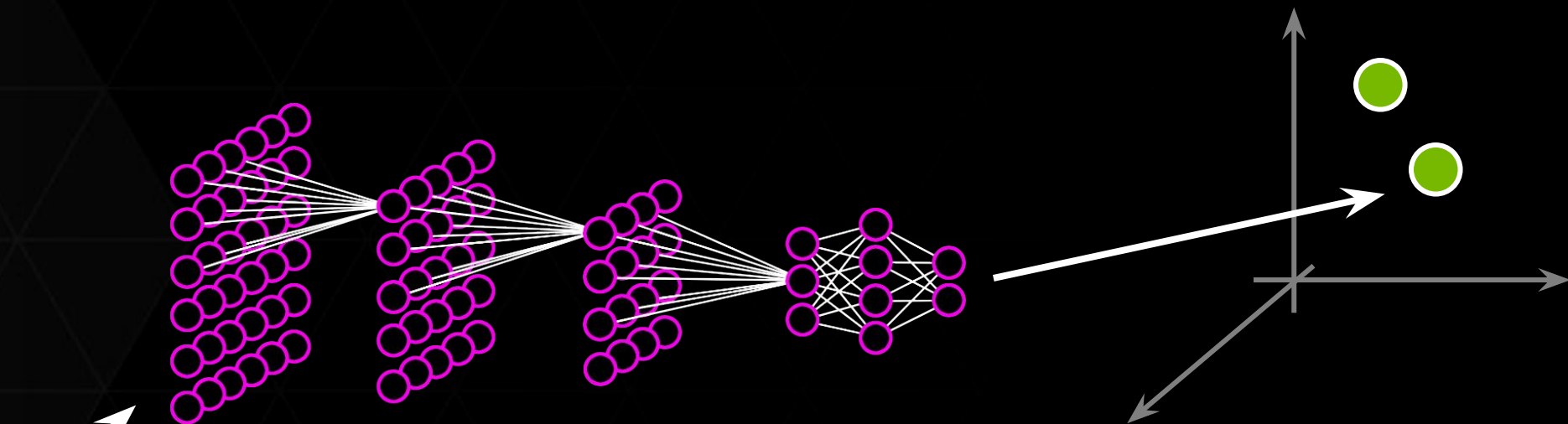




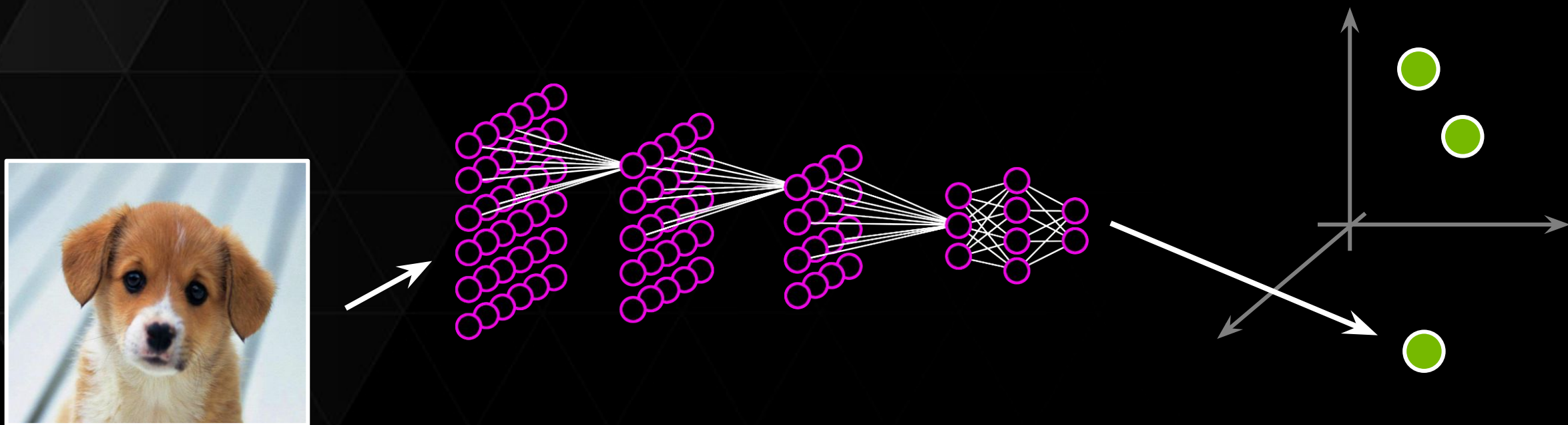
# ОБУЧЕНИЕ ЭМБЕДДИНГОВ



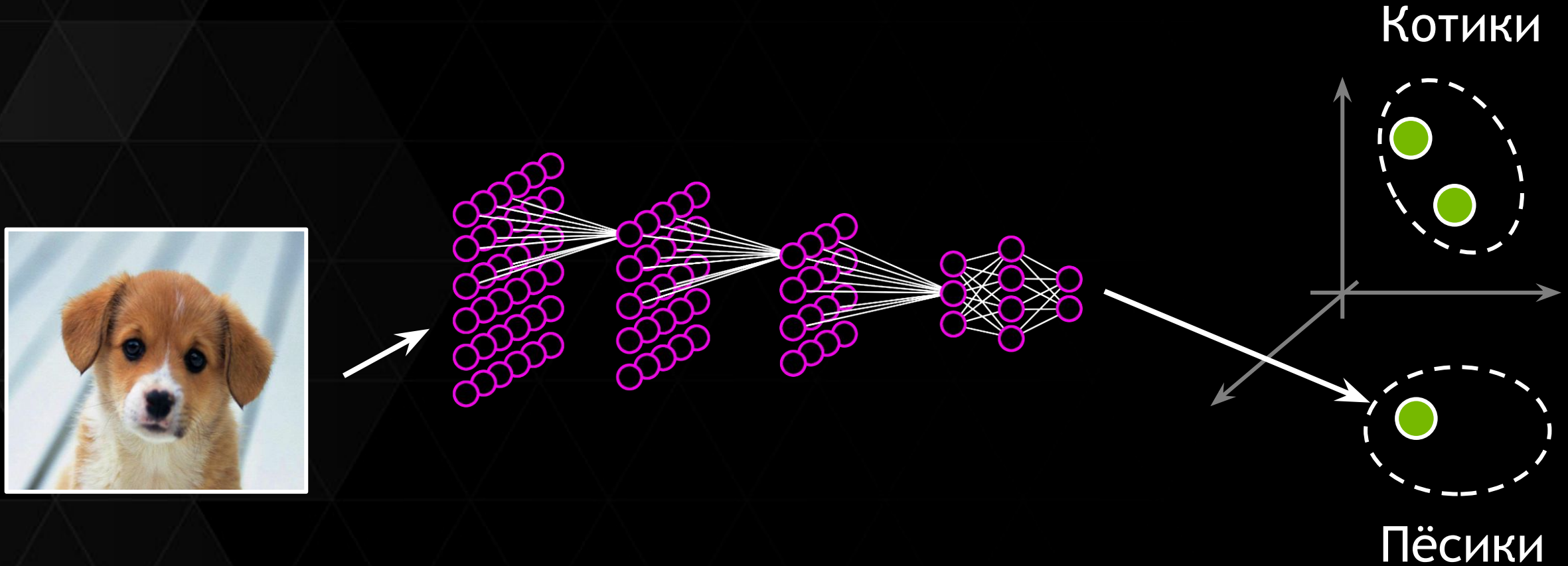
# ОБУЧЕНИЕ ЭМБЕДДИНГОВ



# ОБУЧЕНИЕ ЭМБЕДДИНГОВ

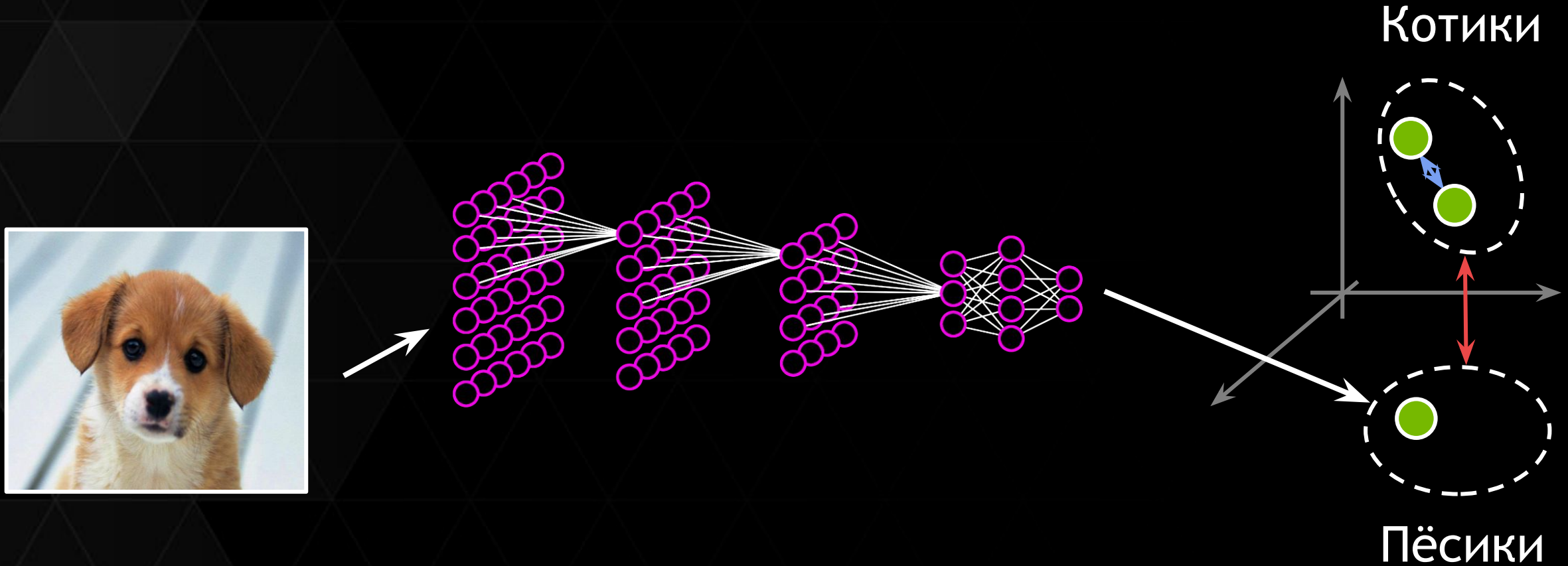


# METRIC LEARNING

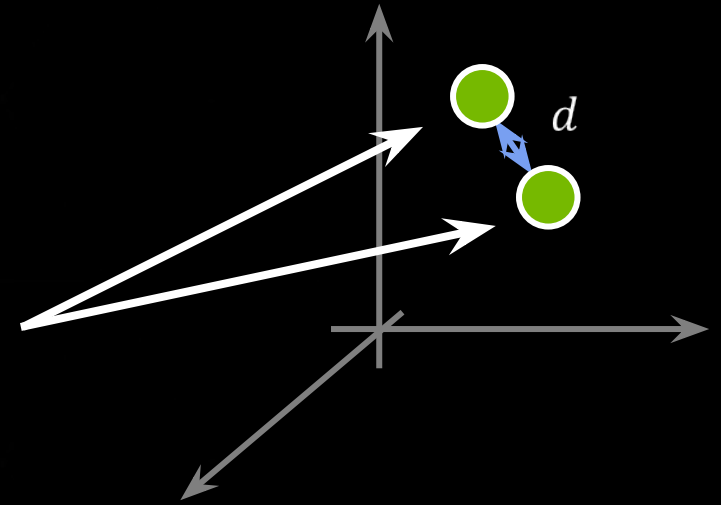
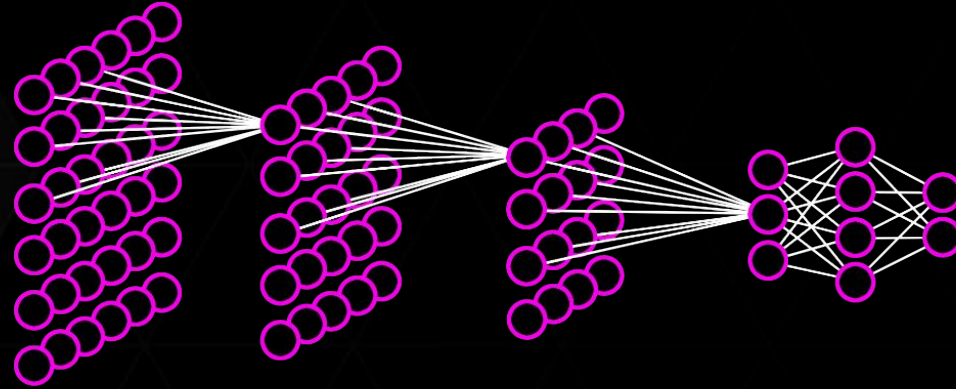




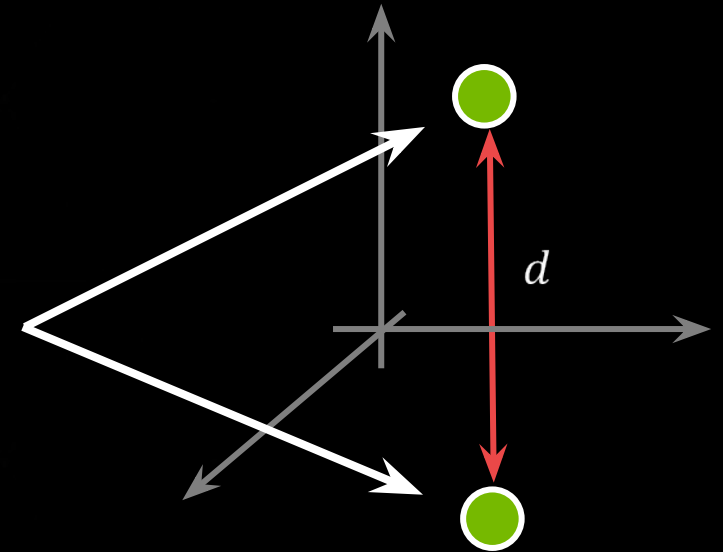
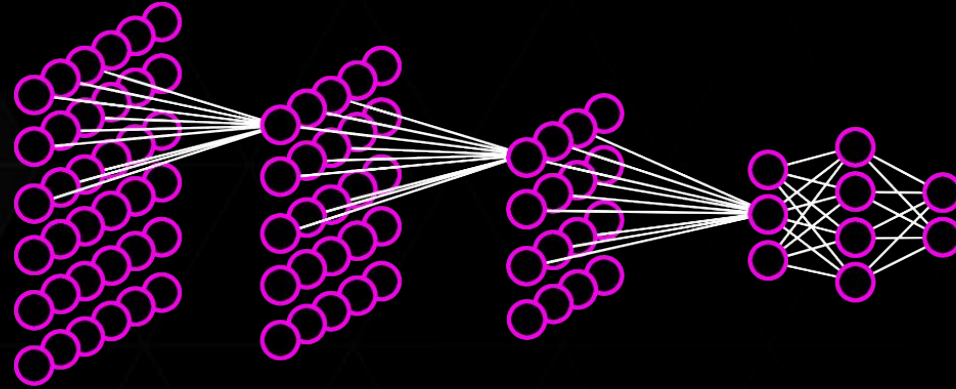
# METRIC LEARNING



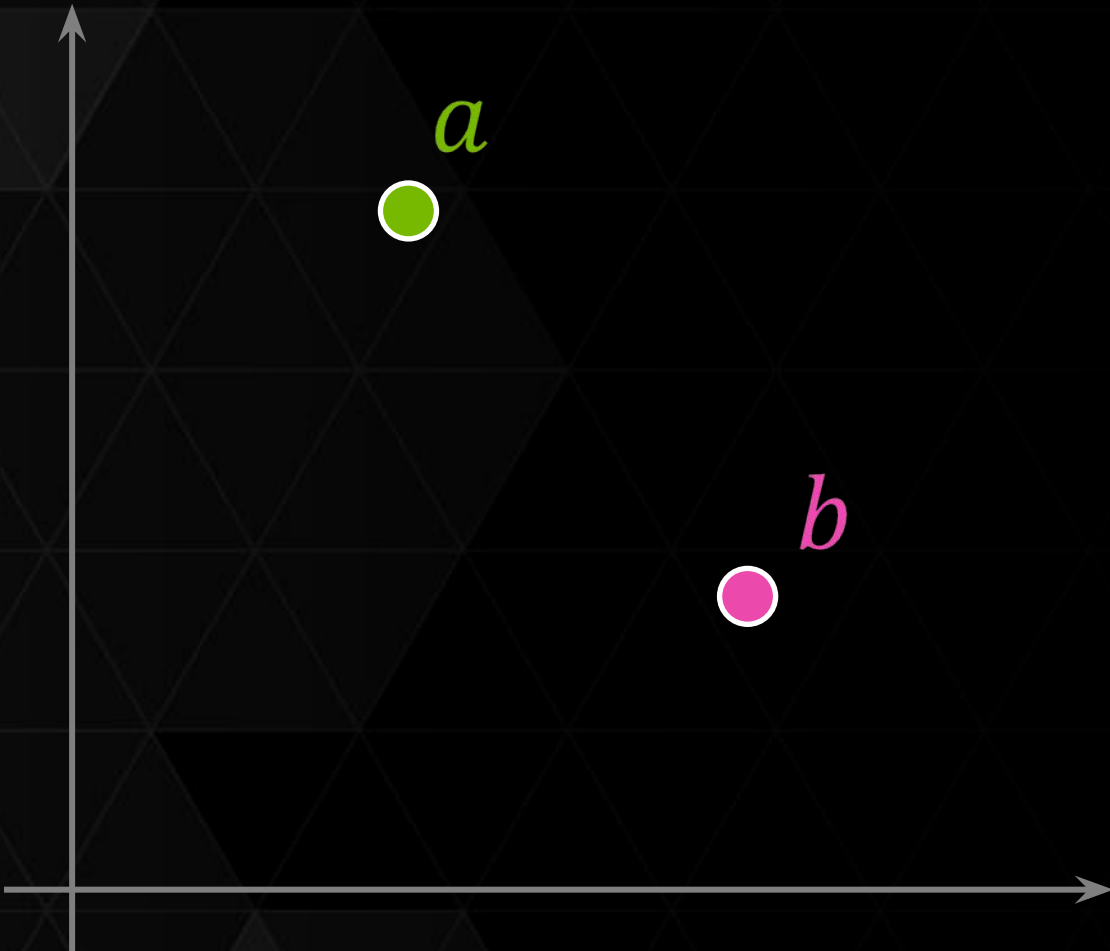
# СИАМСКАЯ НЕЙРОННАЯ СЕТЬ



# СИАМСКАЯ НЕЙРОННАЯ СЕТЬ

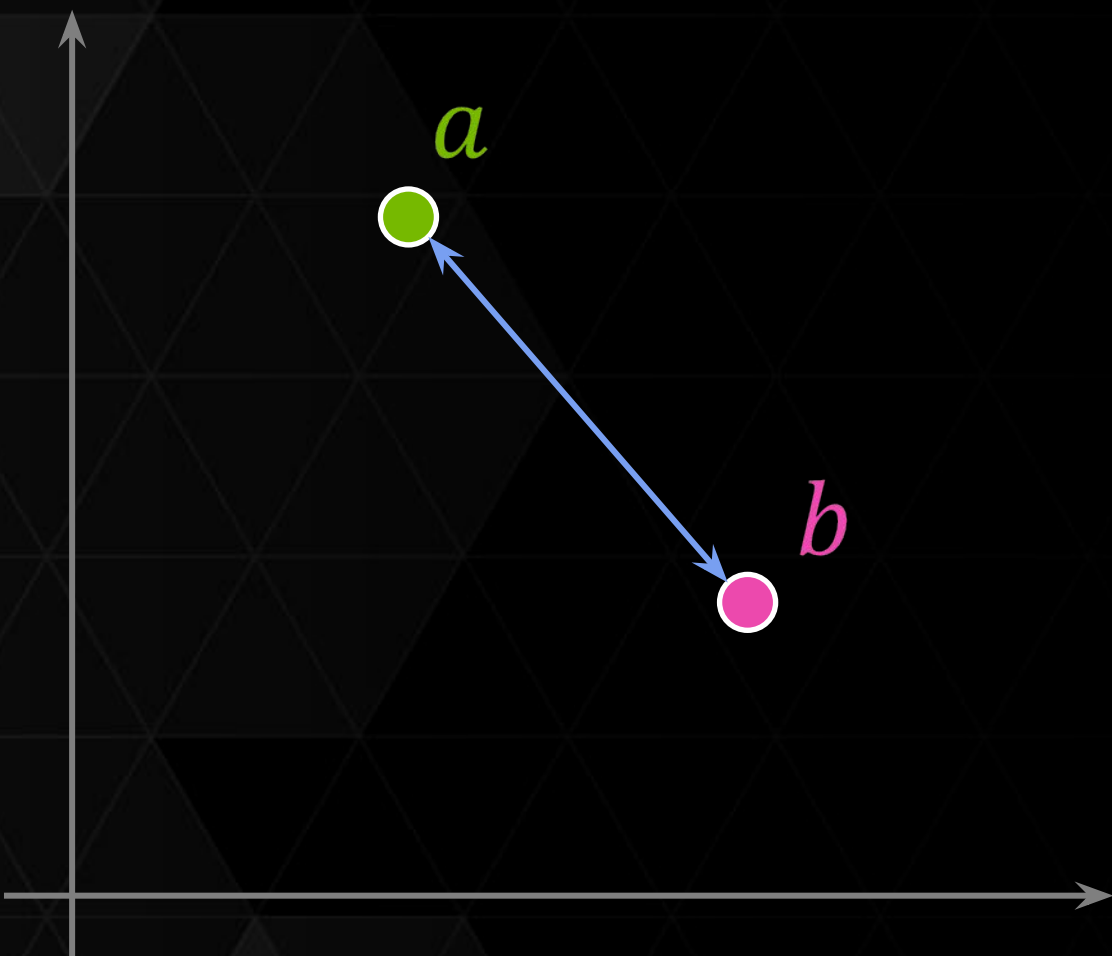


# РАССТОЯНИЕ МЕЖДУ ВЕКТОРАМИ



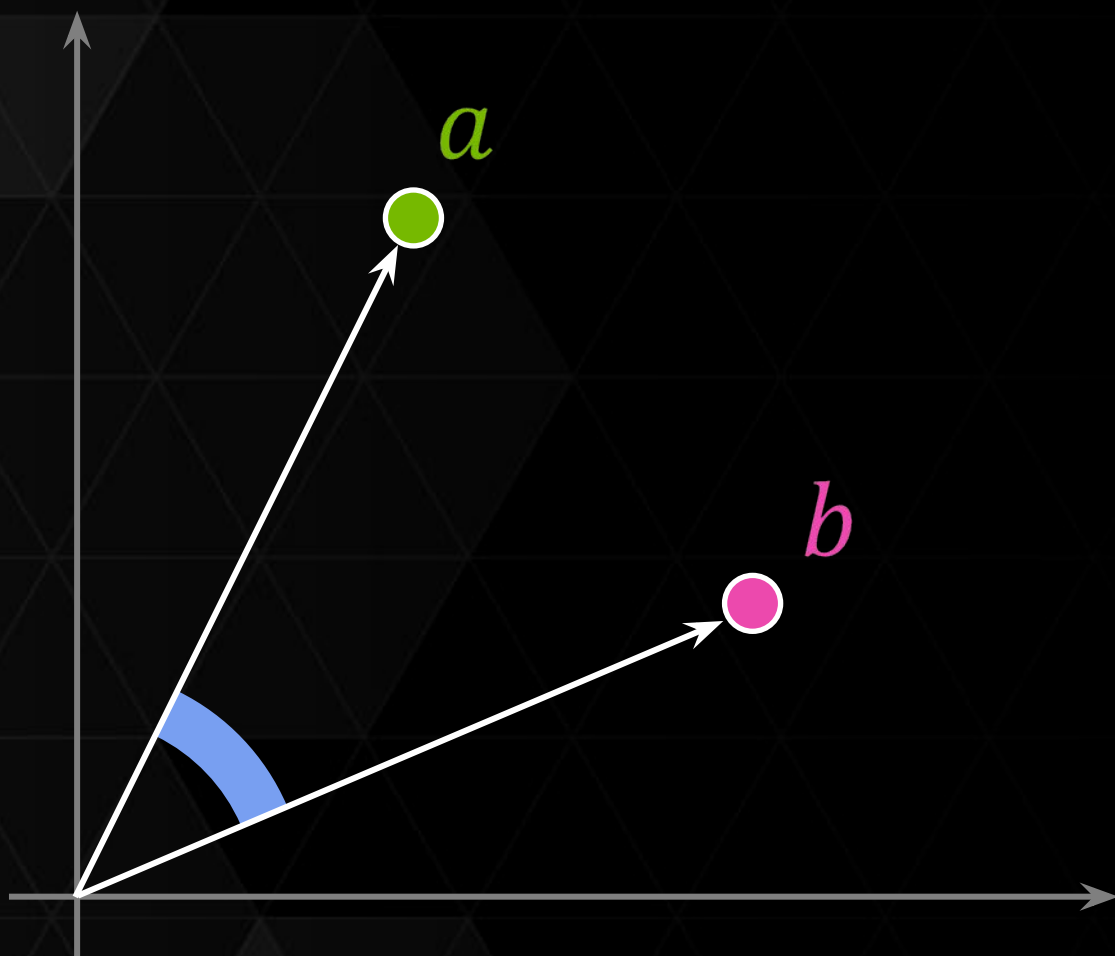


# ЕВКЛИДОВО РАССТОЯНИЕ



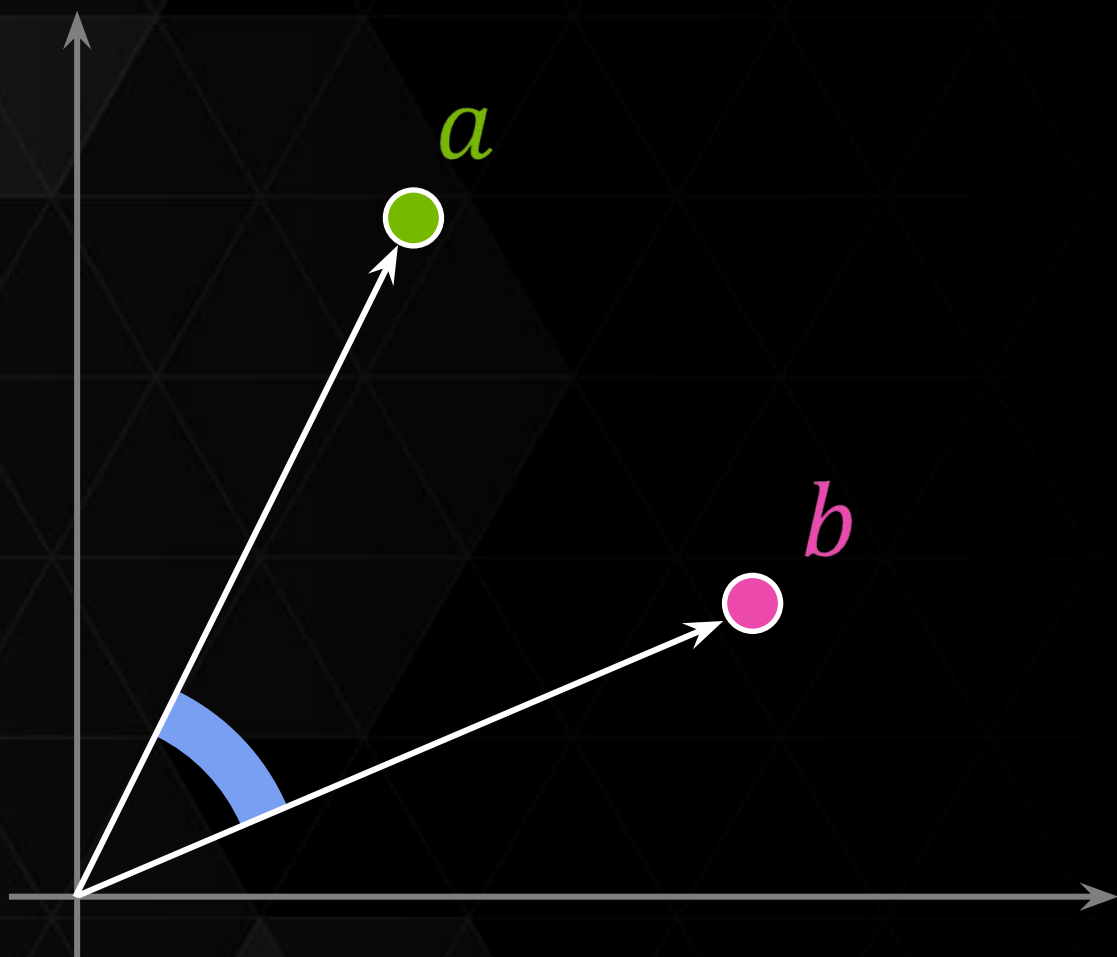
$$D = \|a - b\|$$

# КОСИНУСНОЕ РАССТОЯНИЕ



$$\cos \varphi = \frac{a^T b}{\|a\| * \|b\|}$$

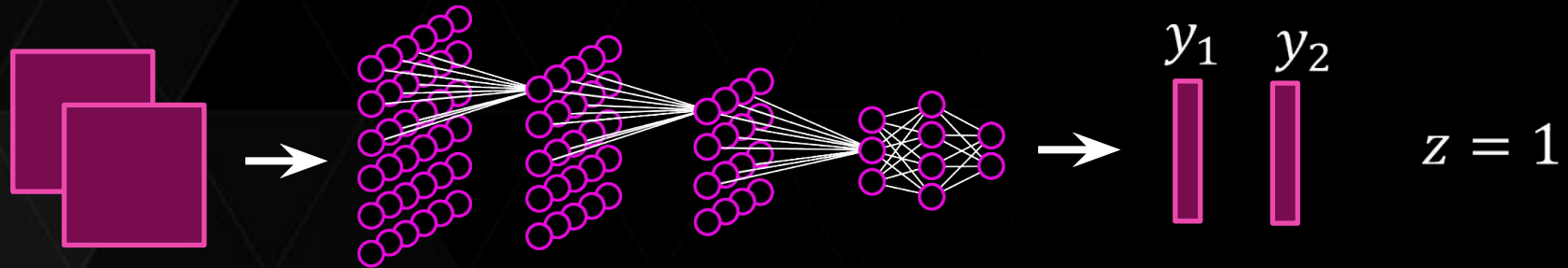
# КОСИНУСНОЕ РАССТОЯНИЕ



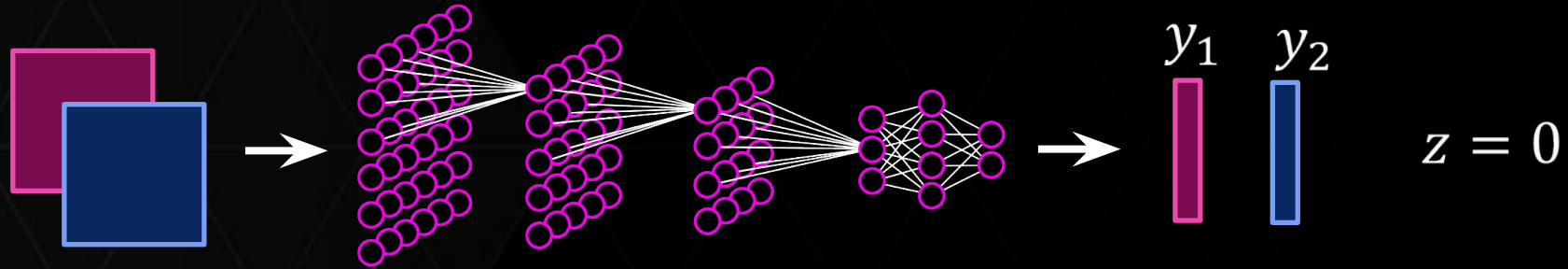
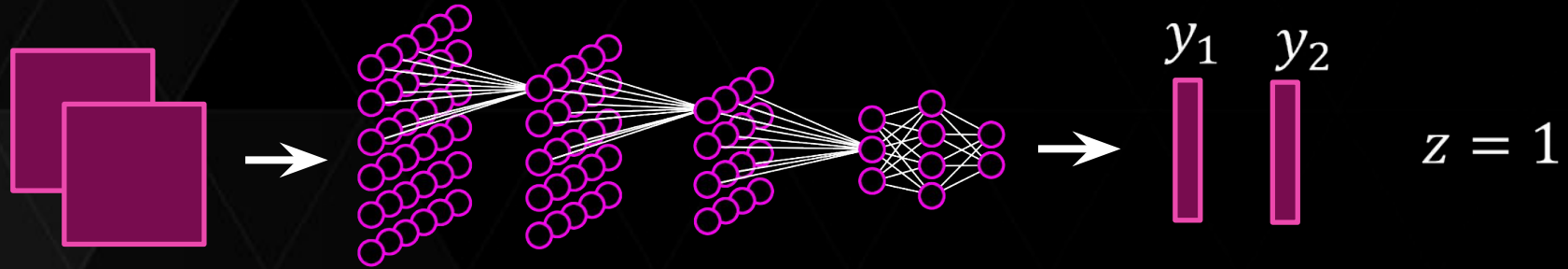
$$\cos \varphi = \frac{a^T b}{\|a\| * \|b\|}$$

$$D = a^T b$$

# СИАМСКАЯ НЕЙРОННАЯ СЕТЬ

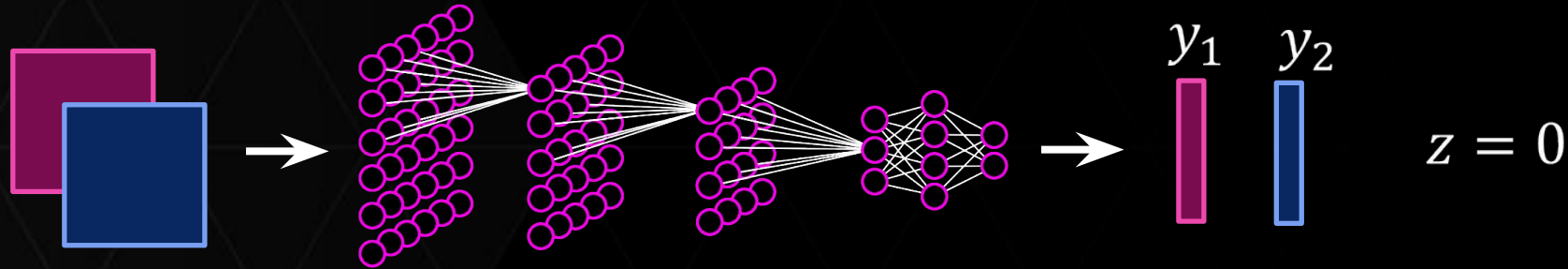
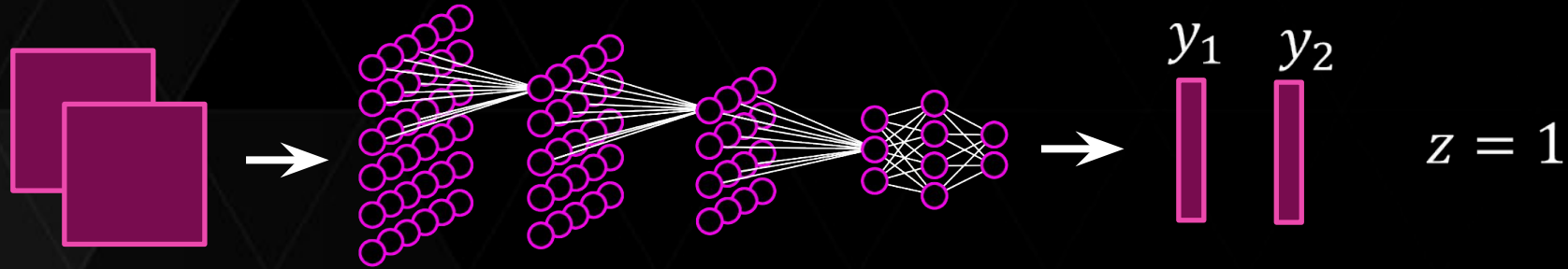


# СИАМСКАЯ НЕЙРОННАЯ СЕТЬ





# СИАМСКАЯ НЕЙРОННАЯ СЕТЬ

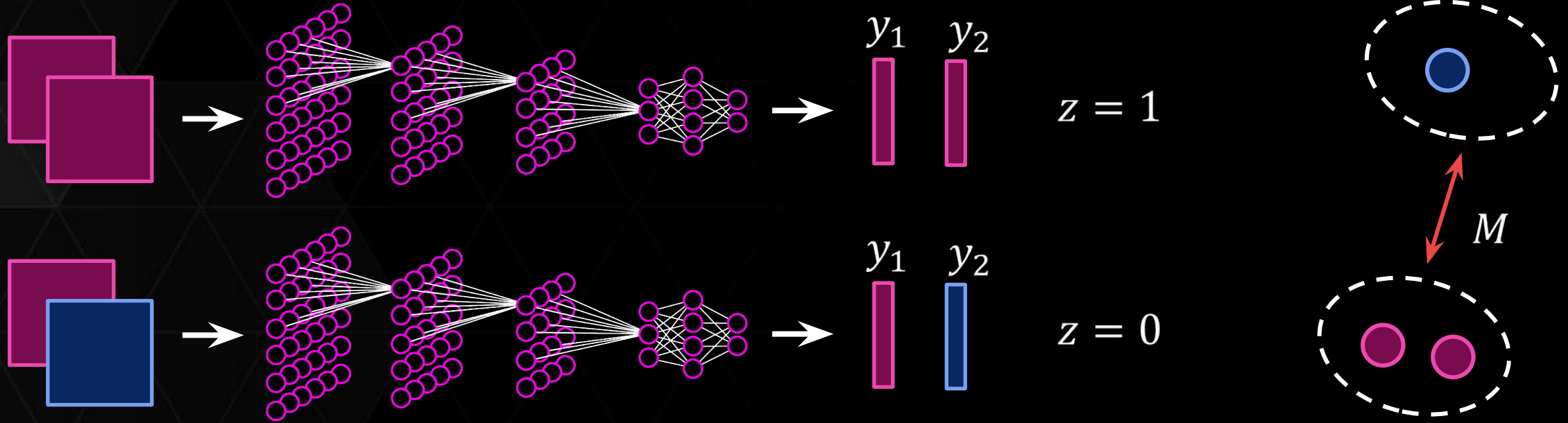


$$d = \|y_1 - y_2\|$$

Contrastive Loss:

$$L = \frac{1}{2} (z * d^2 + (1 - z) * (\max\{0, M - d\})^2)$$

# СИАМСКАЯ НЕЙРОННАЯ СЕТЬ

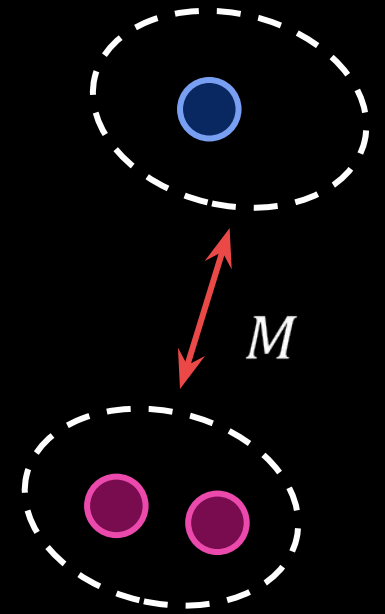
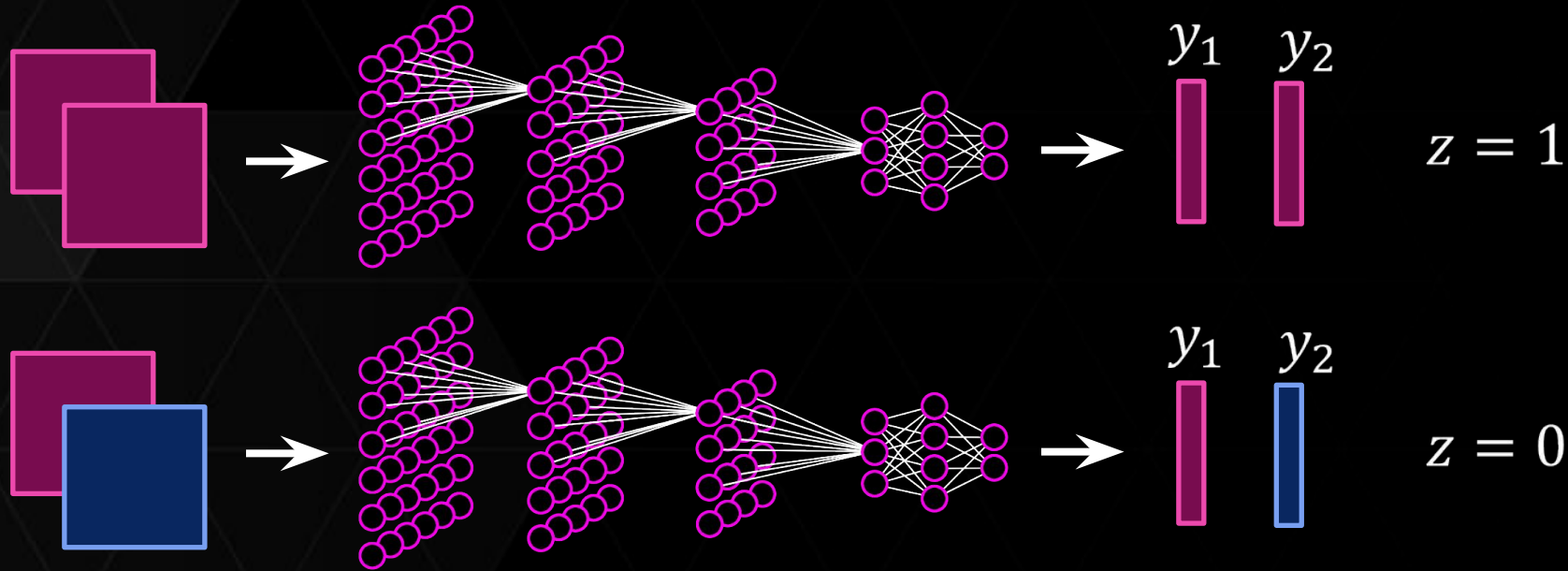


$$d = \|y_1 - y_2\|$$

Contrastive Loss:

$$L = \frac{1}{2} (z * d^2 + (1 - z) * (\max\{0, M - d\})^2)$$

# СИАМСКАЯ НЕЙРОННАЯ СЕТЬ

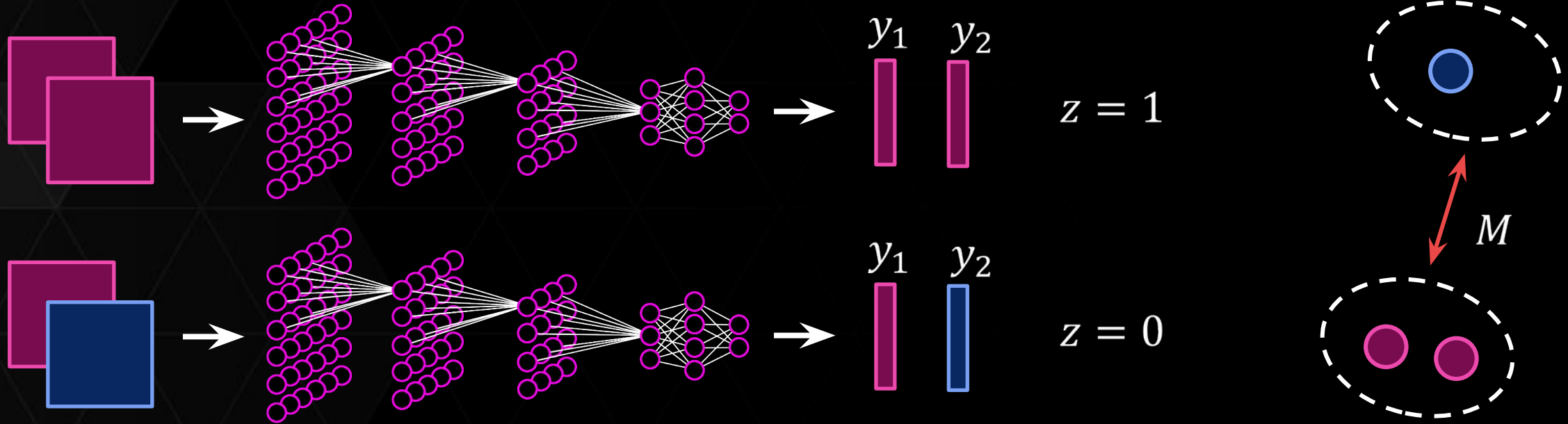


$$d = \|y_1 - y_2\|$$

Contrastive Loss:

$$L = \frac{1}{2} (z * d^2 + (1 - z) * (\max\{0, M - d\})^2)$$

# СИАМСКАЯ НЕЙРОННАЯ СЕТЬ

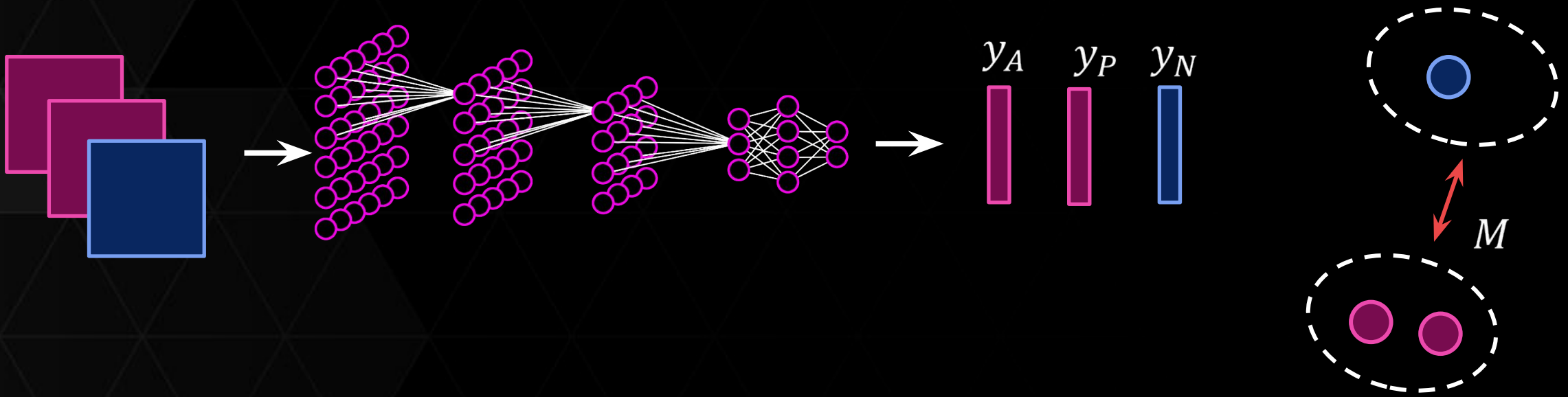


$$d = \|y_1 - y_2\|$$

Contrastive Loss:

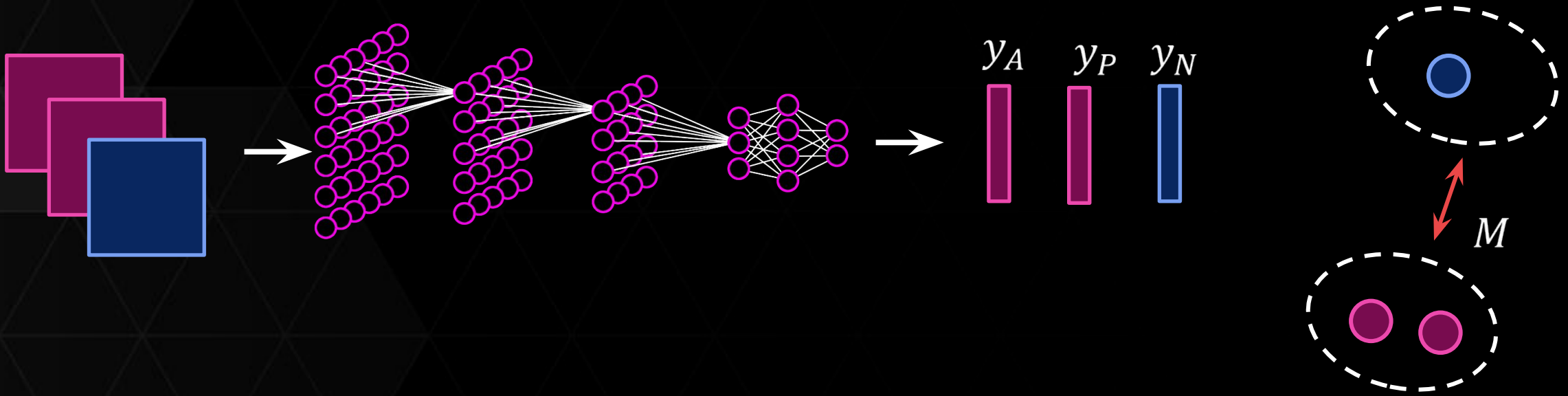
$$L = \frac{1}{2} (z * d^2 + (1 - z) * (\max\{0, M - d\})^2)$$

# TRIPLET LOSS





# TRIPLET LOSS



Triplet Loss: 
$$L = \max\{0, \|y_A - y_P\|^2 - \|y_A - y_N\|^2 + M\}$$

# РАЗЛИЧНЫЕ ЛОССЫ

Softmax Loss

Contrastive Loss

Triplet Loss

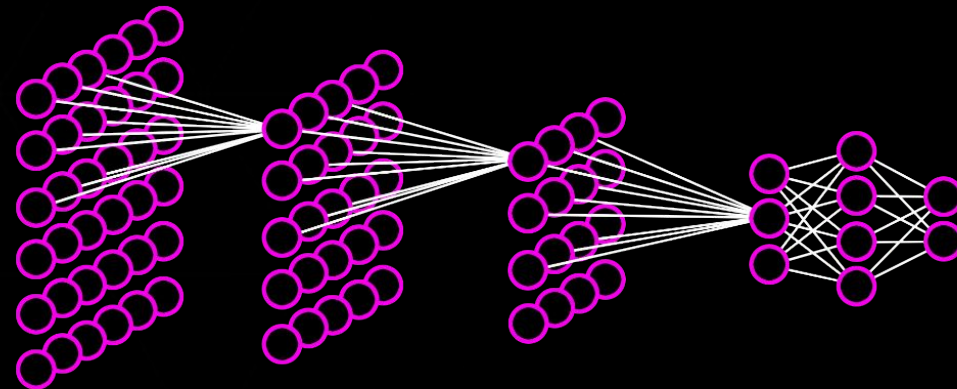
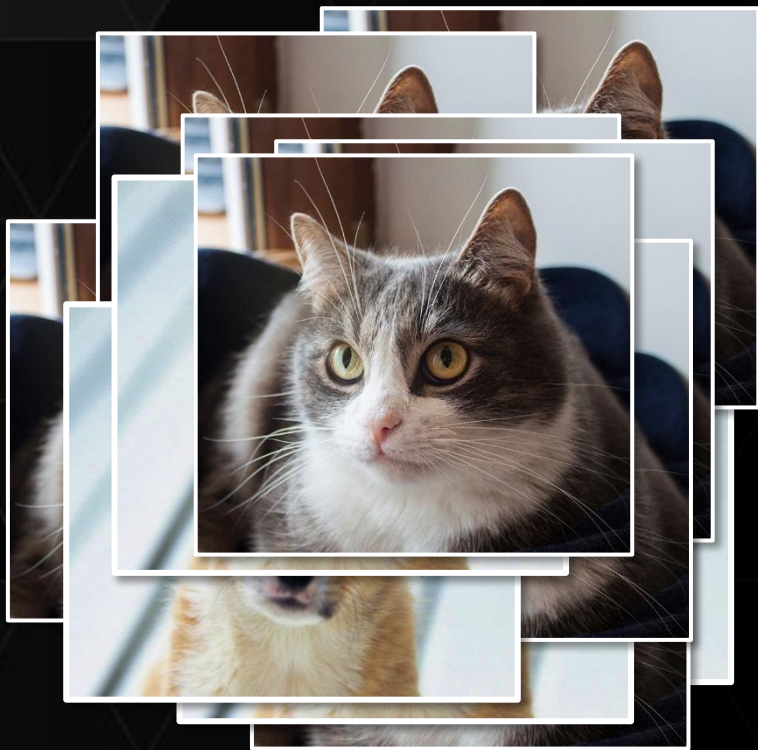
Center Loss

SphereFace

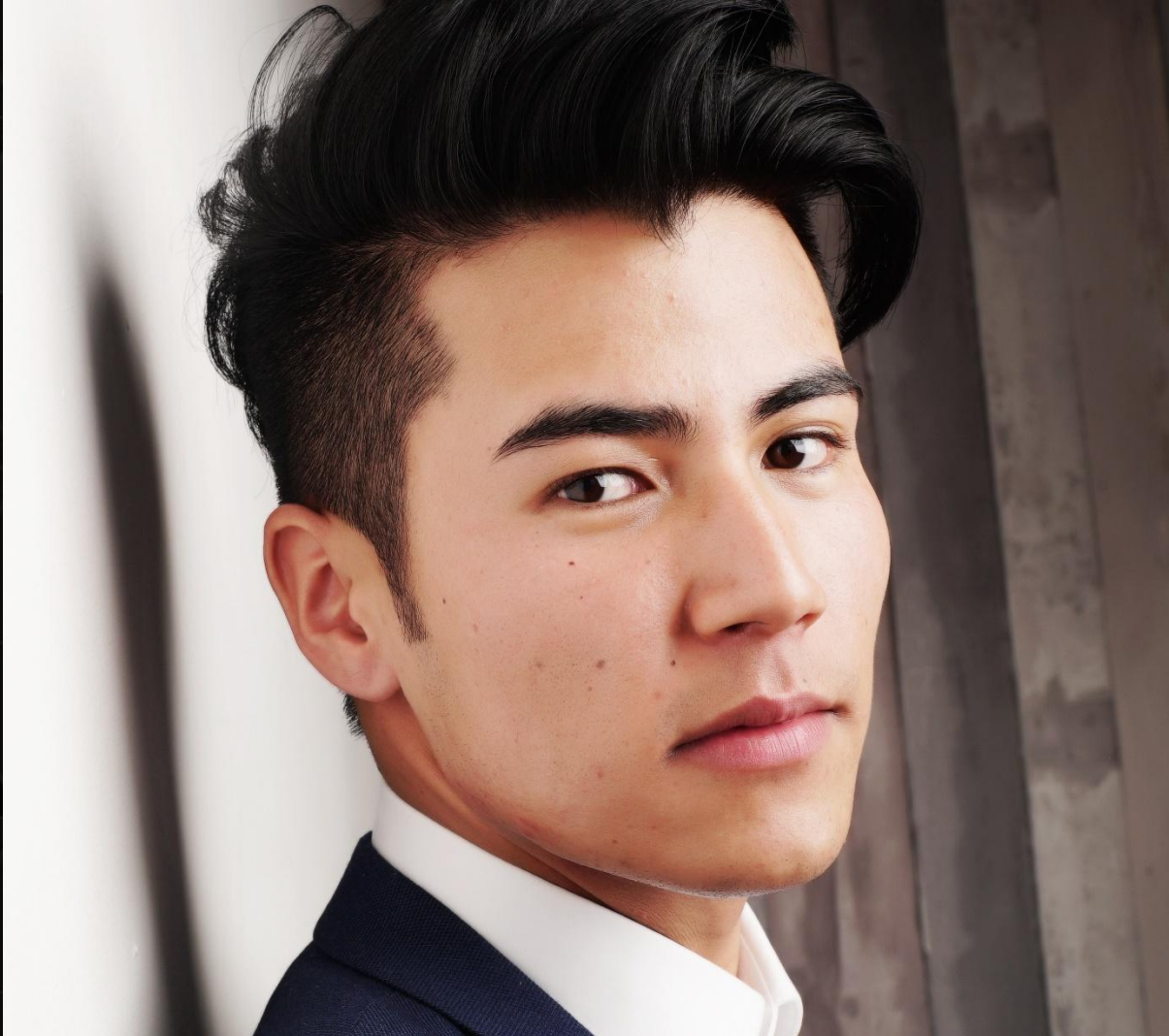
ArcFace

# ПРОБЛЕМА: НУЖНО МНОГО ДАННЫХ

Обучающая выборка

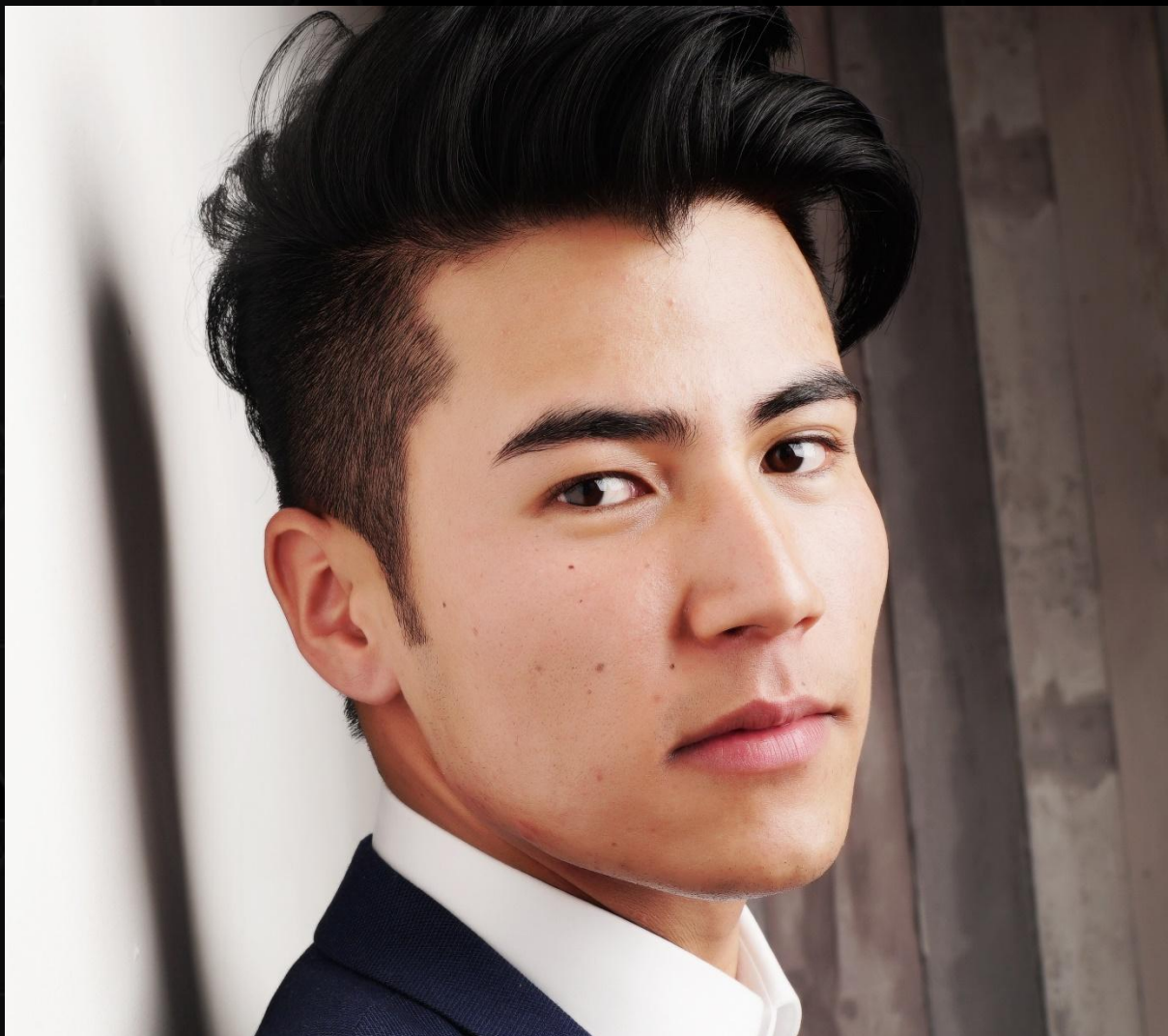


# ПРОБЛЕМА: НУЖНО МНОГО ДАННЫХ





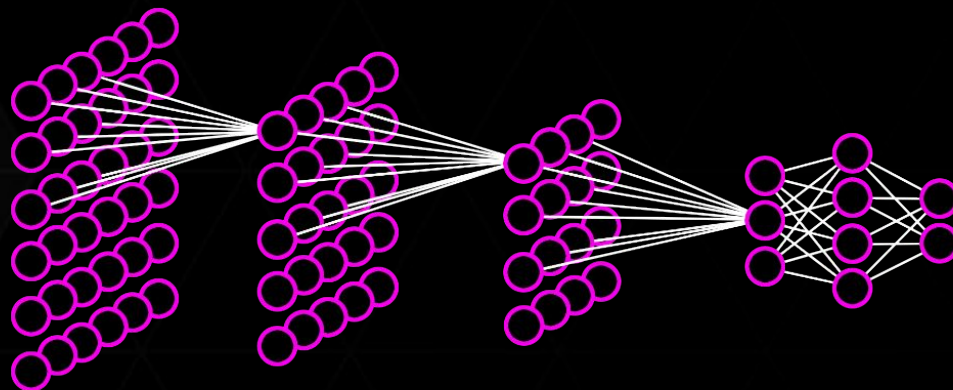
# ПРОБЛЕМА: НУЖНО МНОГО ДАННЫХ





# ONE-SHOT LEARNING

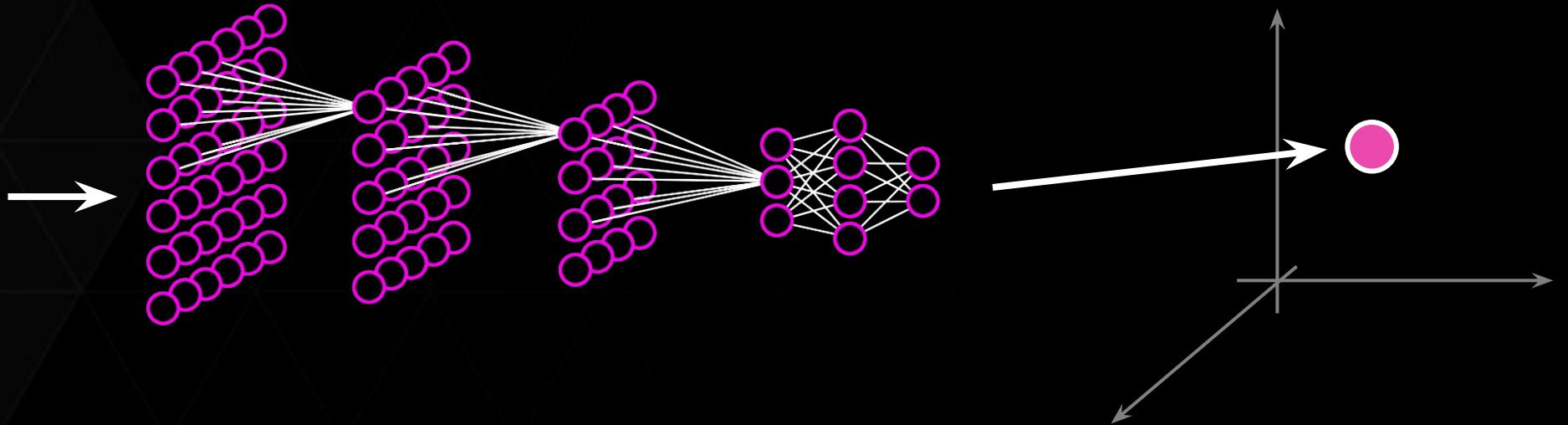
Модель, обученная делать «хорошие» эмбединги



# ONE-SHOT LEARNING

«Обучение»

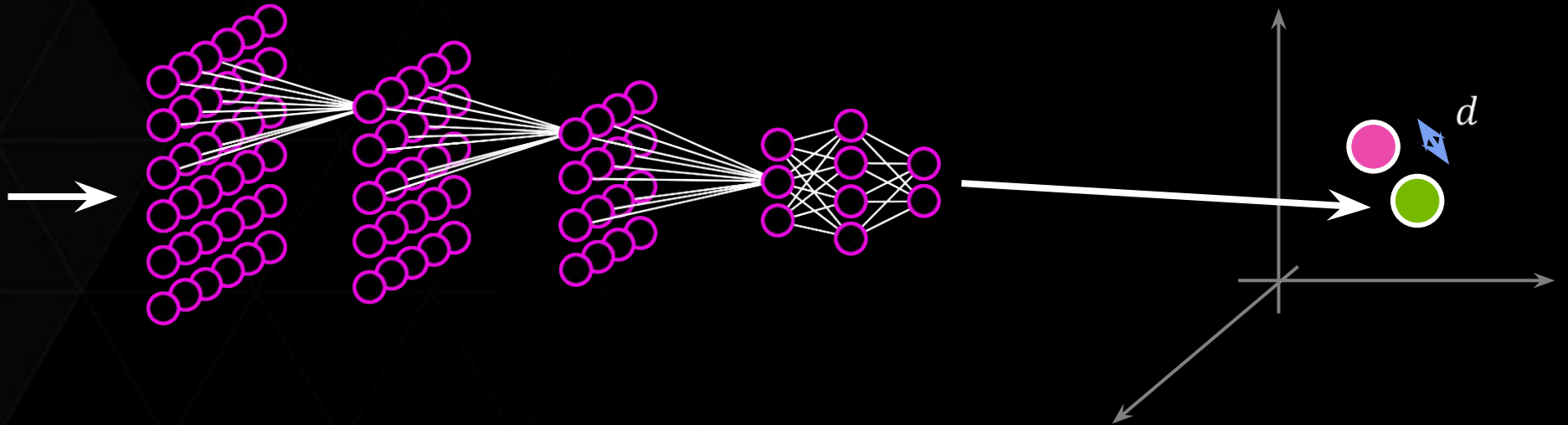
Новый класс



# ONE-SHOT LEARNING

«Инференс»

Новый объект



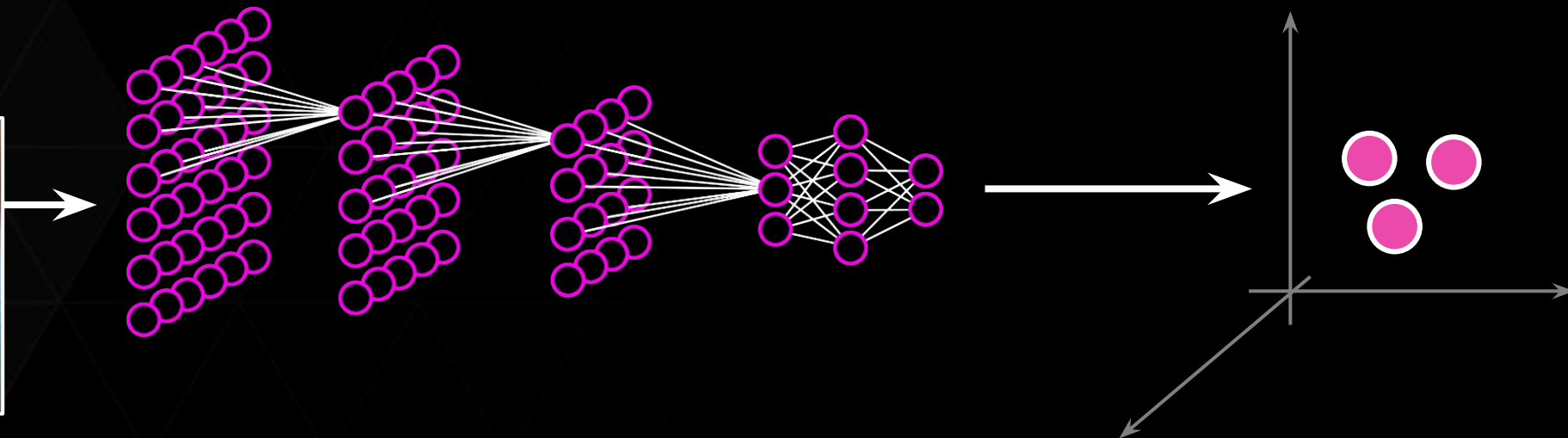
# FEW-SHOT LEARNING

«Обучение»

Новый класс



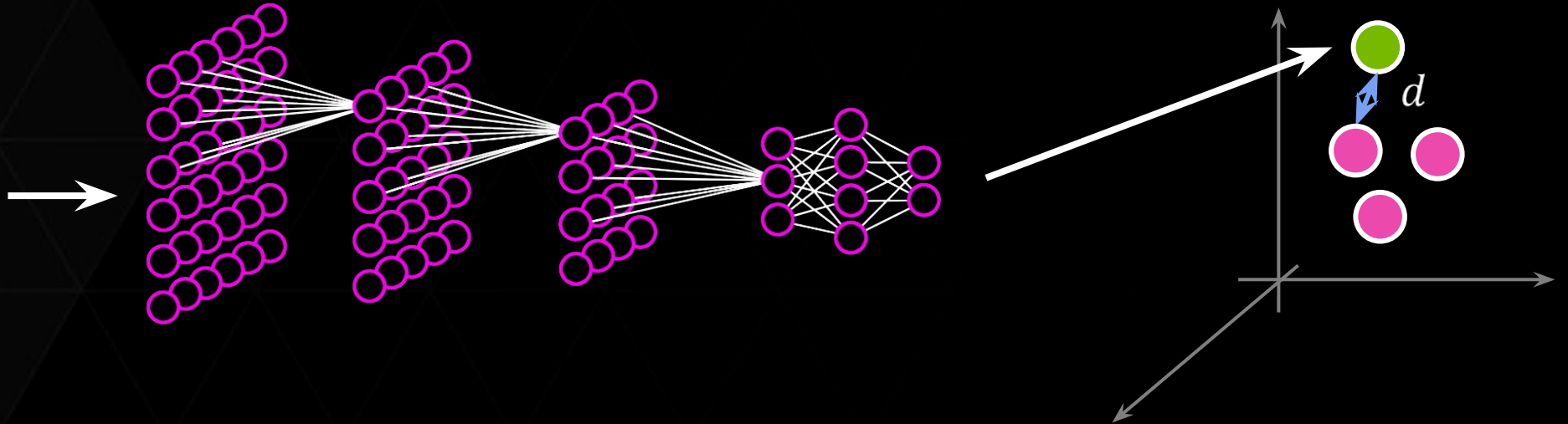
Несколько  
экземпляров



# FEW-SHOT LEARNING

«Инференс»

Новый объект





# ZERO-SHOT LEARNING

«Обучение»

Новый класс

Информация  
о классе



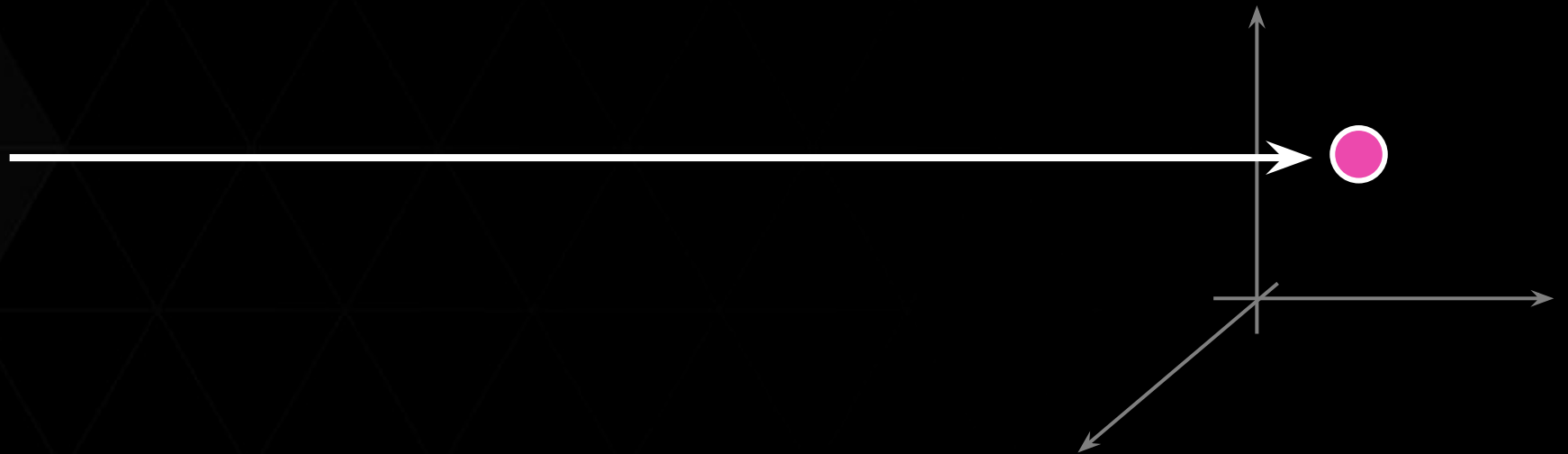
# ZERO-SHOT LEARNING

«Обучение»

Новый класс



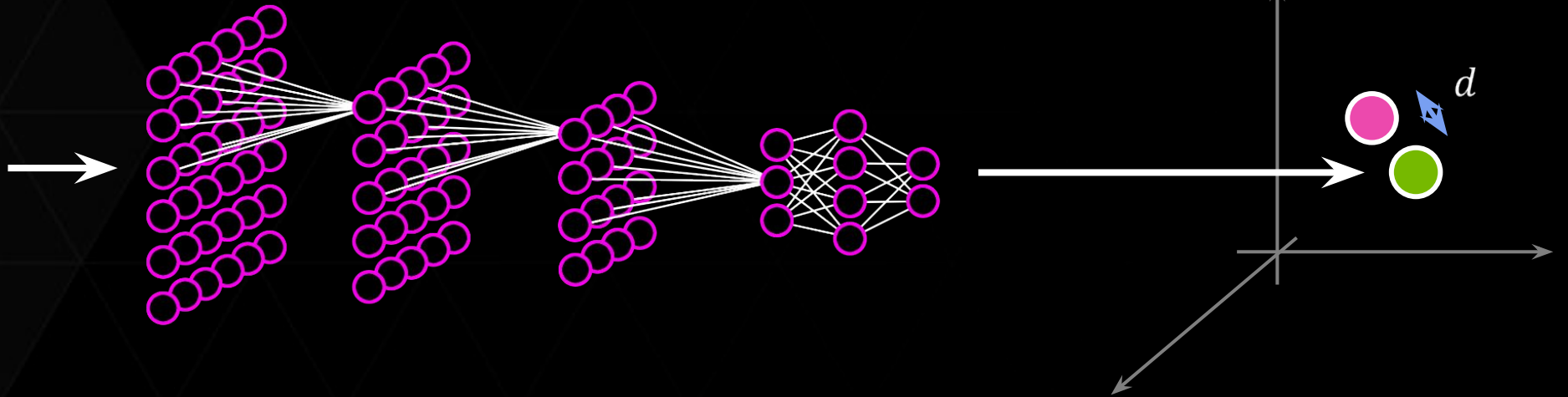
Прототип класса



# ZERO-SHOT LEARNING

«Инференс»

Новый объект



# ПОИСК БЛИЖАЙШИХ СОСЕДЕЙ

Матрица всех  
эмбедингов

$X$

Евклидово  
расстояние

$$\|x_i - y\|$$

# ПОИСК БЛИЖАЙШИХ СОСЕДЕЙ

Матрица всех  
эмбедингов

$X$

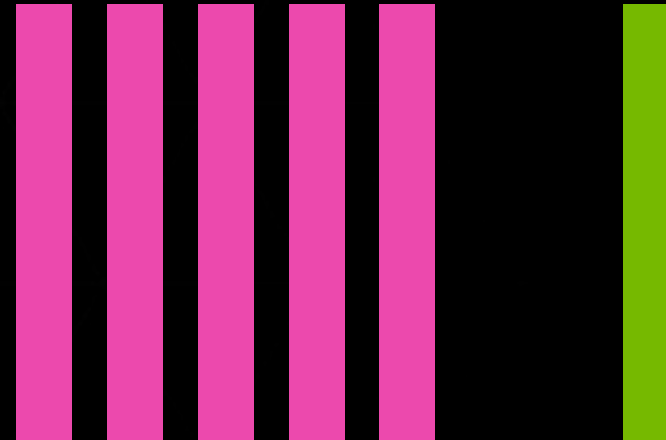
Евклидово  
расстояние

$$\|x_i - y\|$$

Поиск ближайшего соседа

$X$

$y$





# ПОИСК БЛИЖАЙШИХ СОСЕДЕЙ

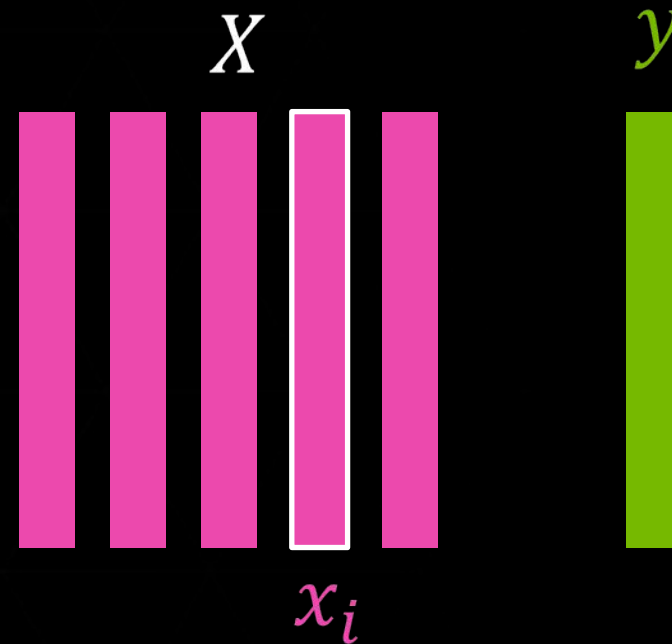
Матрица всех  
эмбеддингов

$X$

Евклидово  
расстояние

$$\|x_i - y\|$$

Поиск ближайшего соседа



# ПОИСК БЛИЖАЙШИХ СОСЕДЕЙ

Матрица всех  
эмбедингов

$X$

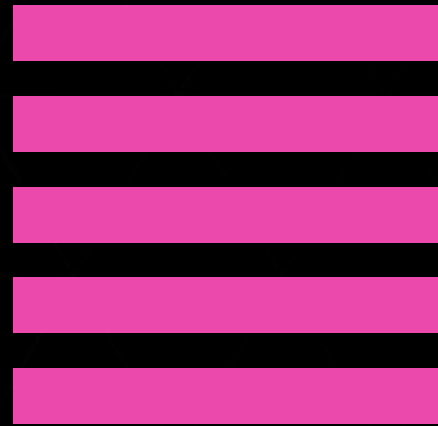
Косинусное  
расстояние

$x_i^T y$

Поиск ближайшего соседа

$X^T$

$y$



\*



# ПОИСК БЛИЖАЙШИХ СОСЕДЕЙ

Матрица всех  
эмбедингов

$X$

Косинусное  
расстояние

$x_i^T y$

Поиск ближайшего соседа

$$\begin{array}{c} X^T \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{array} * \begin{array}{c} y \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{array} = \begin{array}{c} X^T y \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{array}$$

# ПОИСК БЛИЖАЙШИХ СОСЕДЕЙ

Матрица всех  
эмбеддингов

$X$

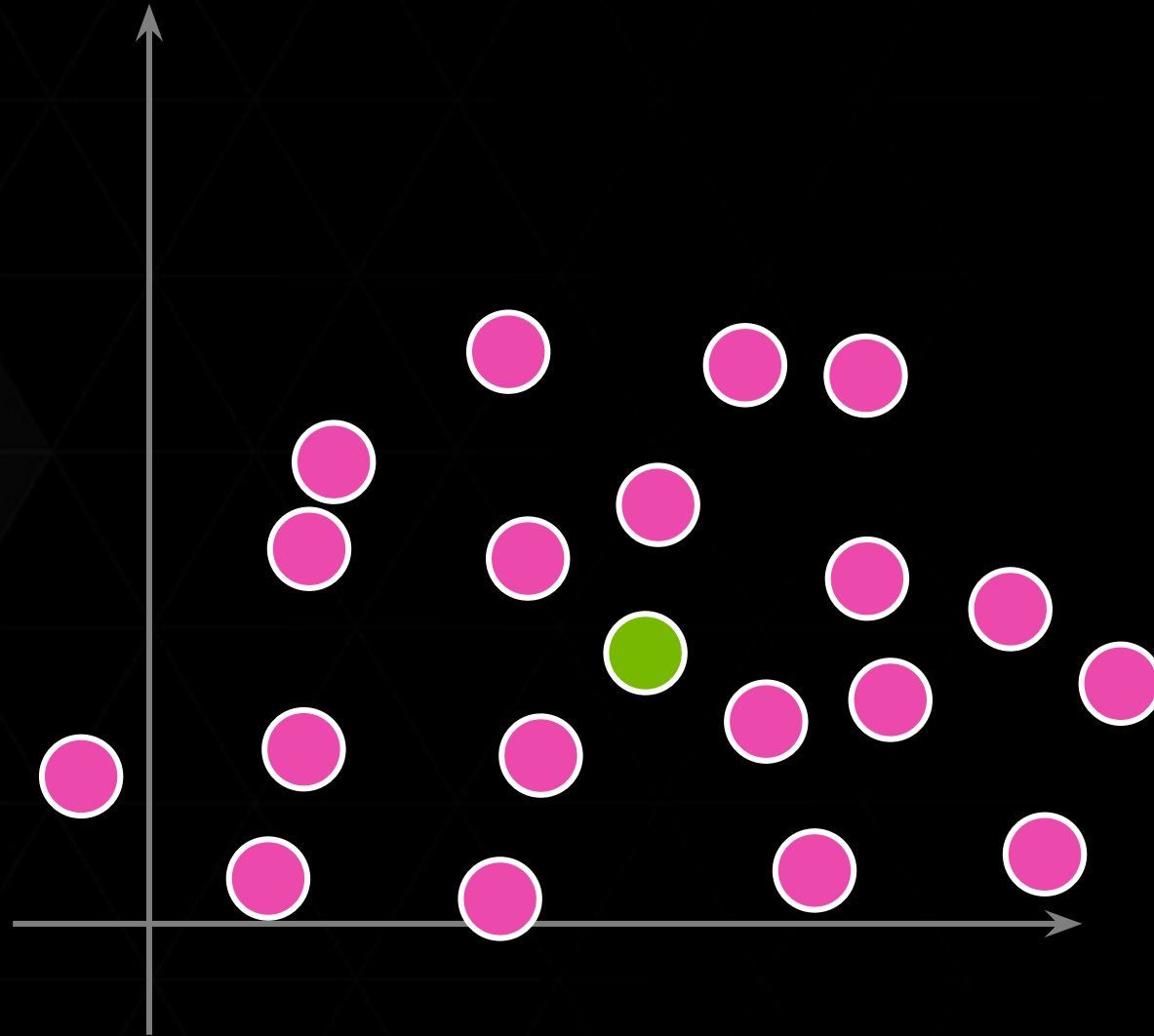
Косинусное  
расстояние

$x_i^T y$

Поиск ближайшего соседа

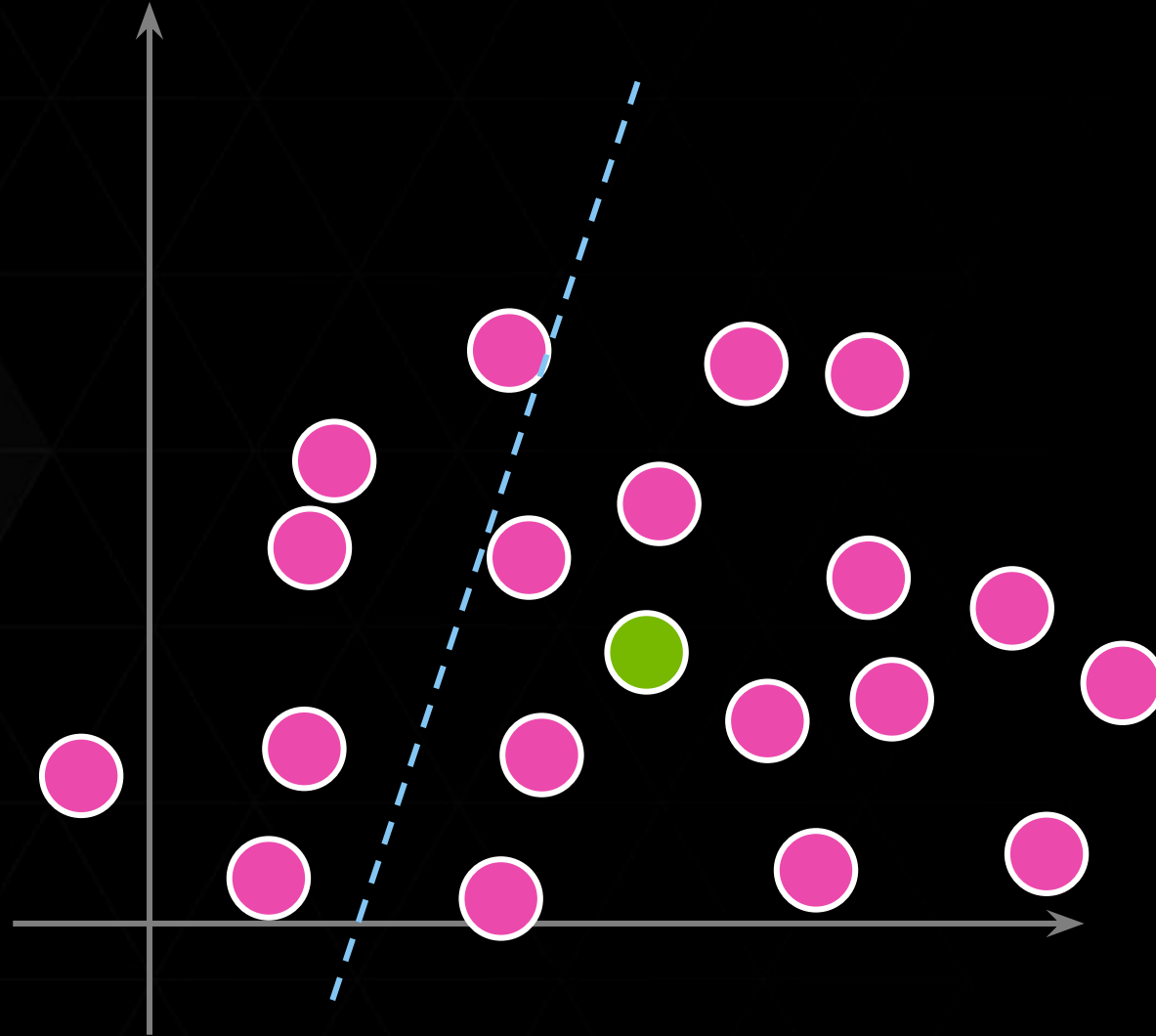
$$\begin{array}{c} x_i \\ \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{array} \end{array} \overset{X^T}{\text{---}} * \begin{array}{c} y \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{array} = \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{array} \overset{X^T y}{\text{---}} \quad x_i^T y$$

# LOCALITY-SENSITIVE HASH (LSH)



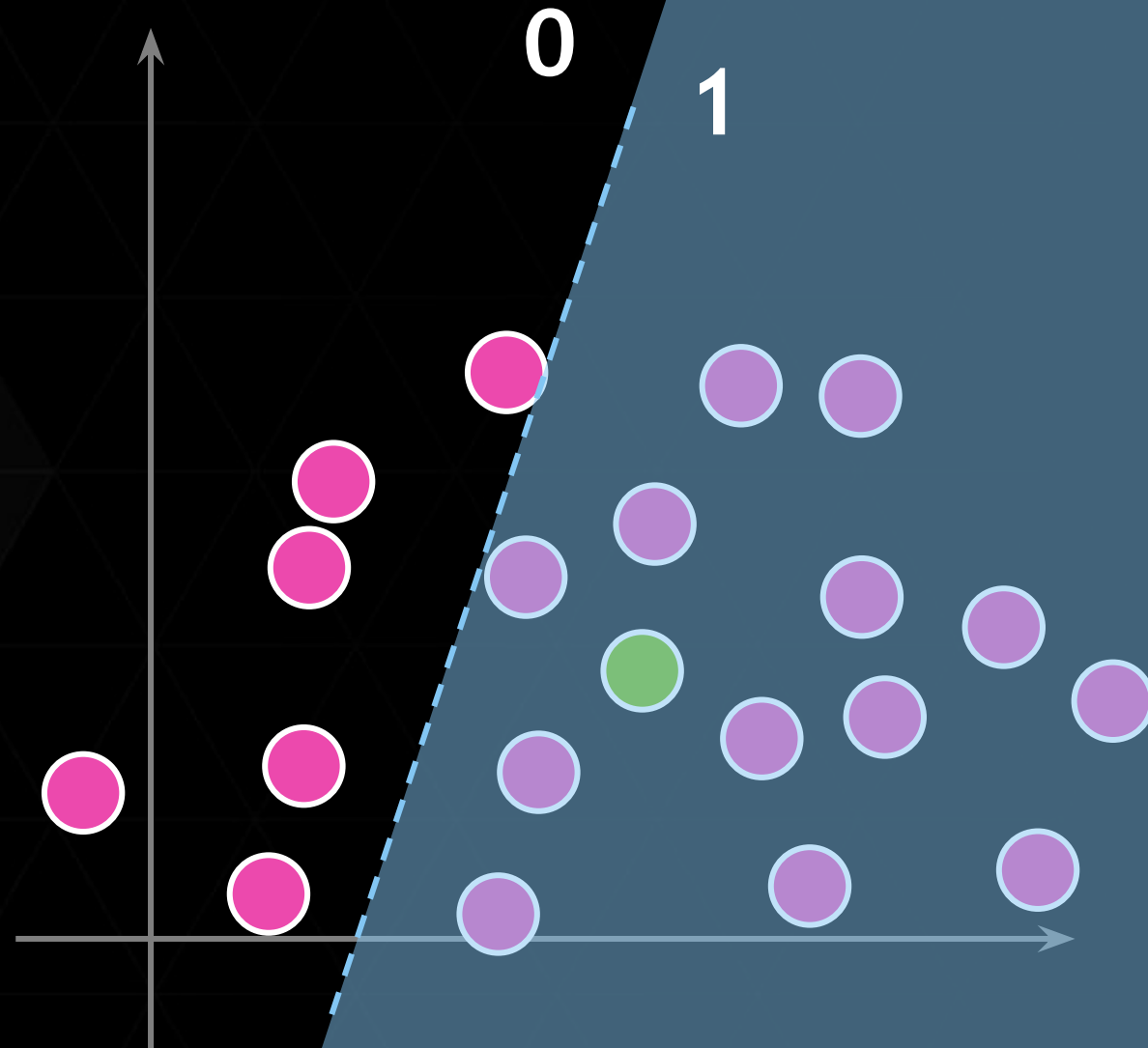


# LOCALITY-SENSITIVE HASH (LSH)



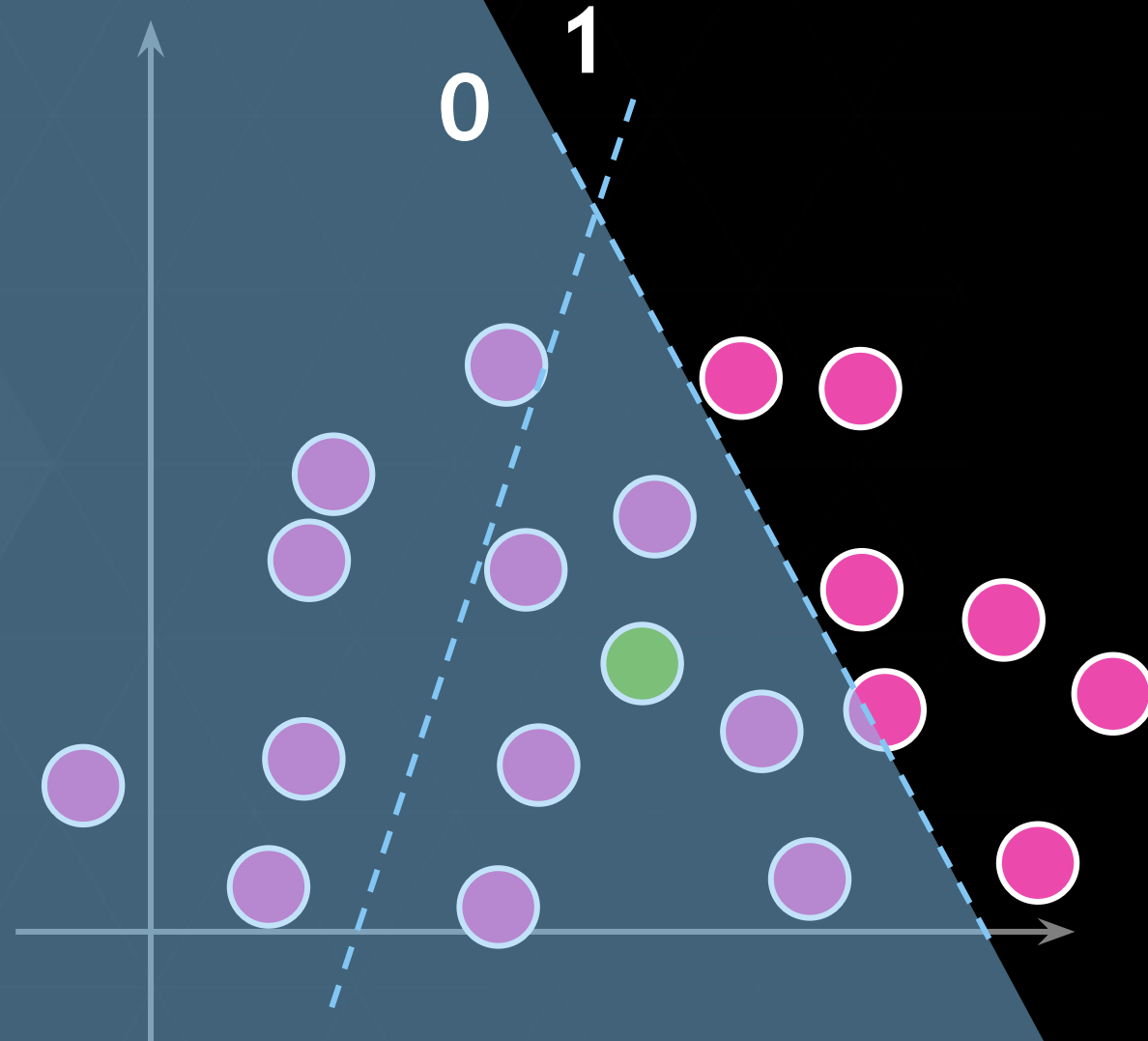
# LOCALITY-SENSITIVE HASH (LSH)

Hash: 1



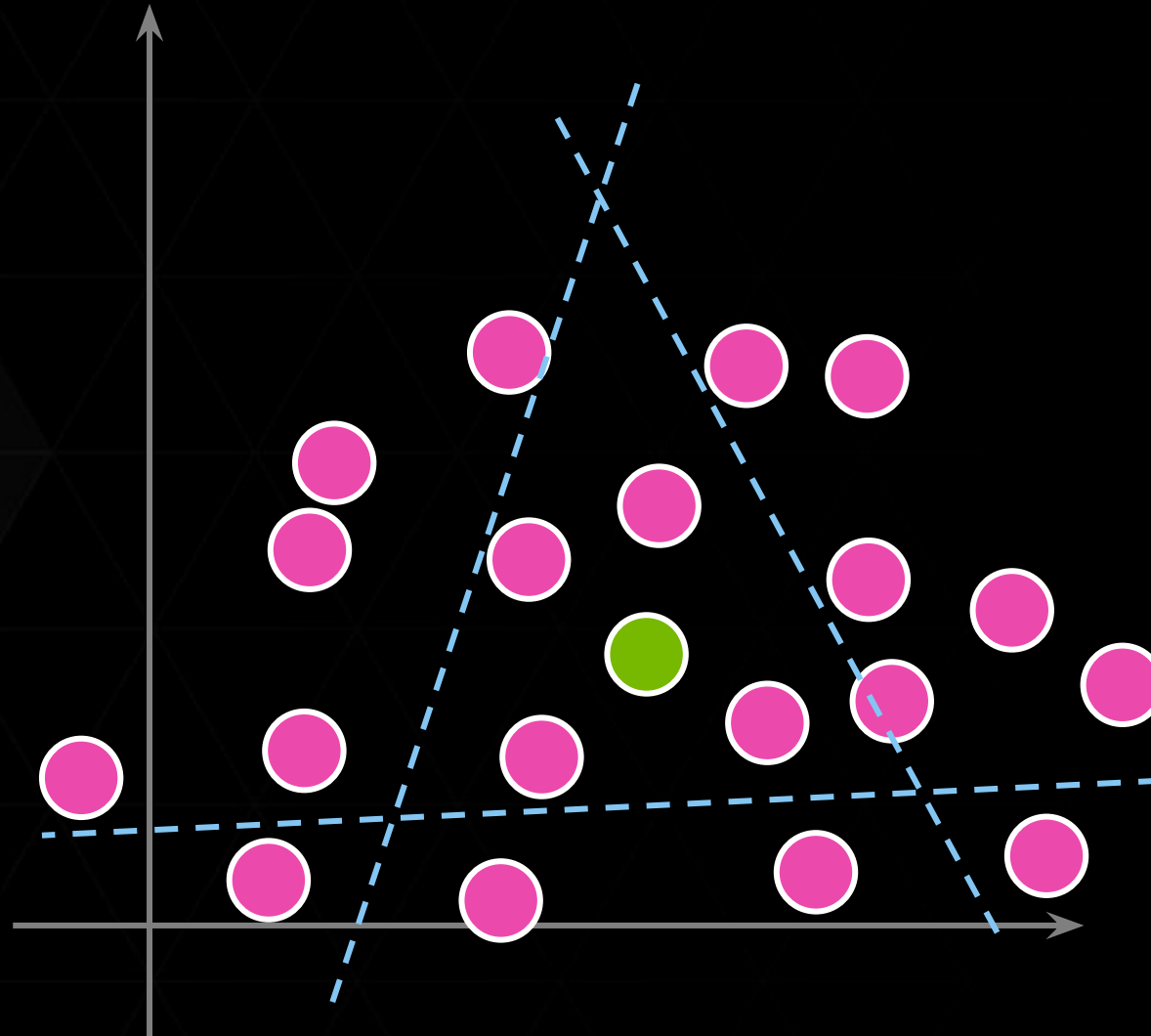
# LOCALITY-SENSITIVE HASH (LSH)

Hash: 10



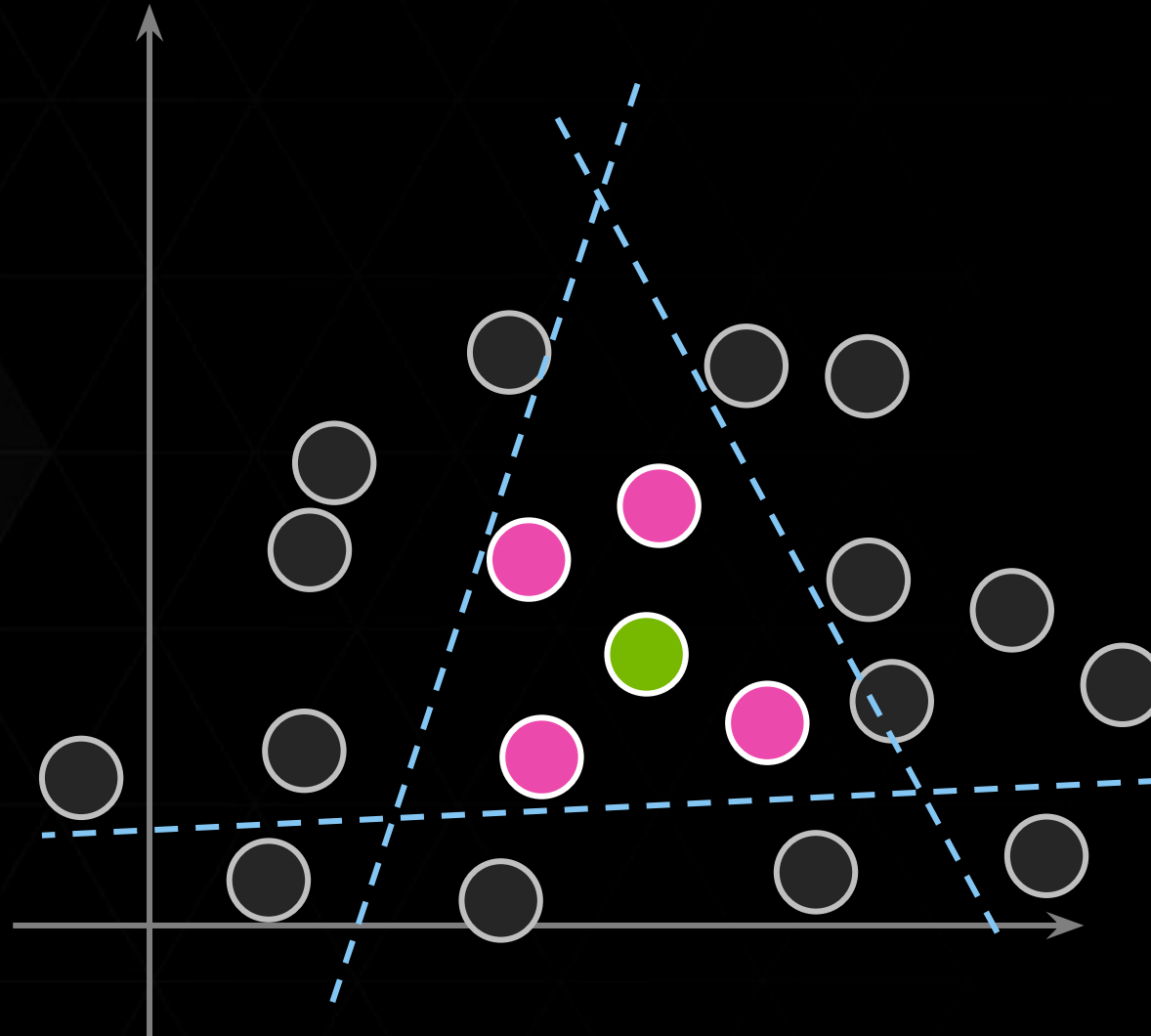
# LOCALITY-SENSITIVE HASH (LSH)

Hash: 101...



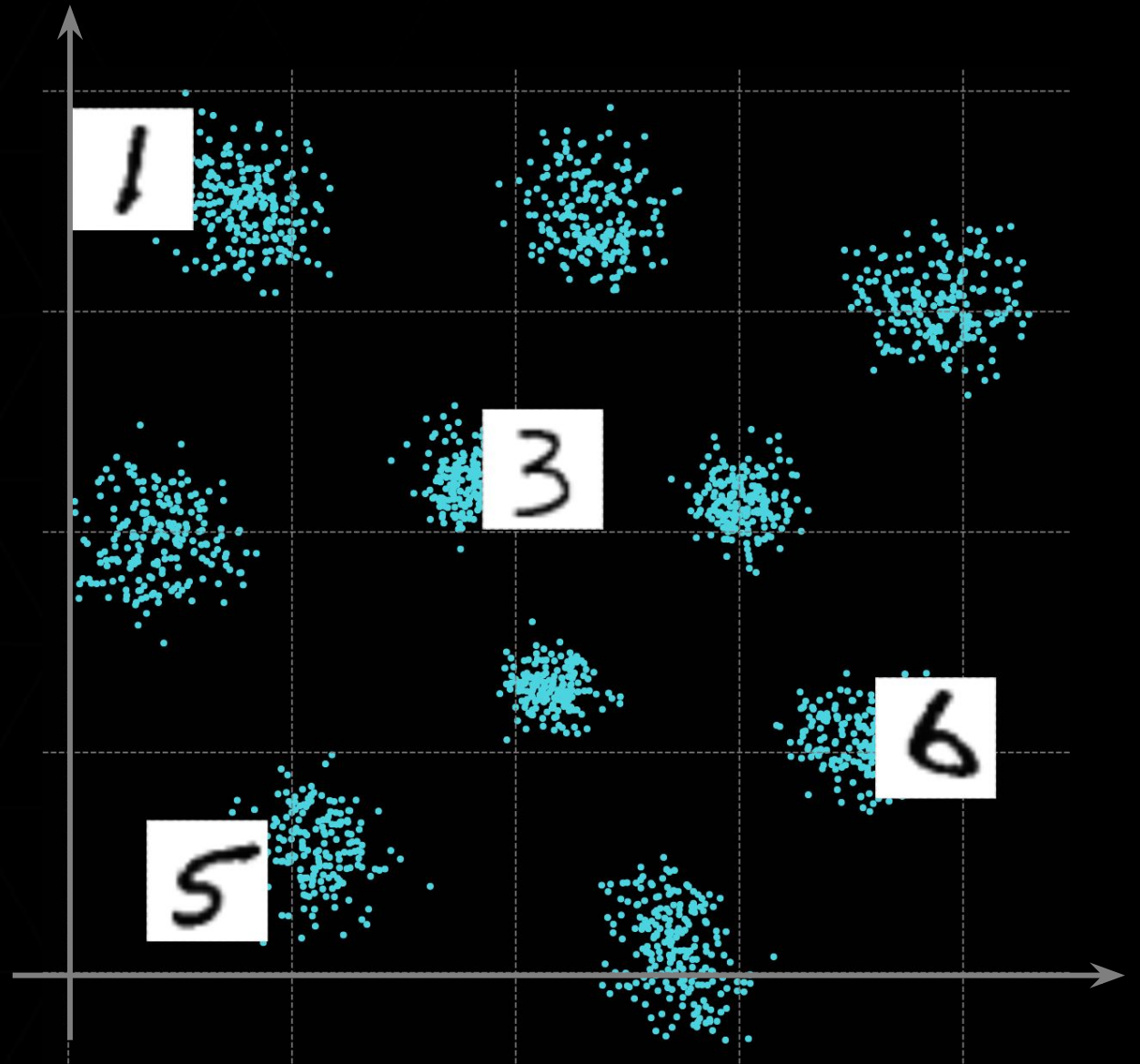
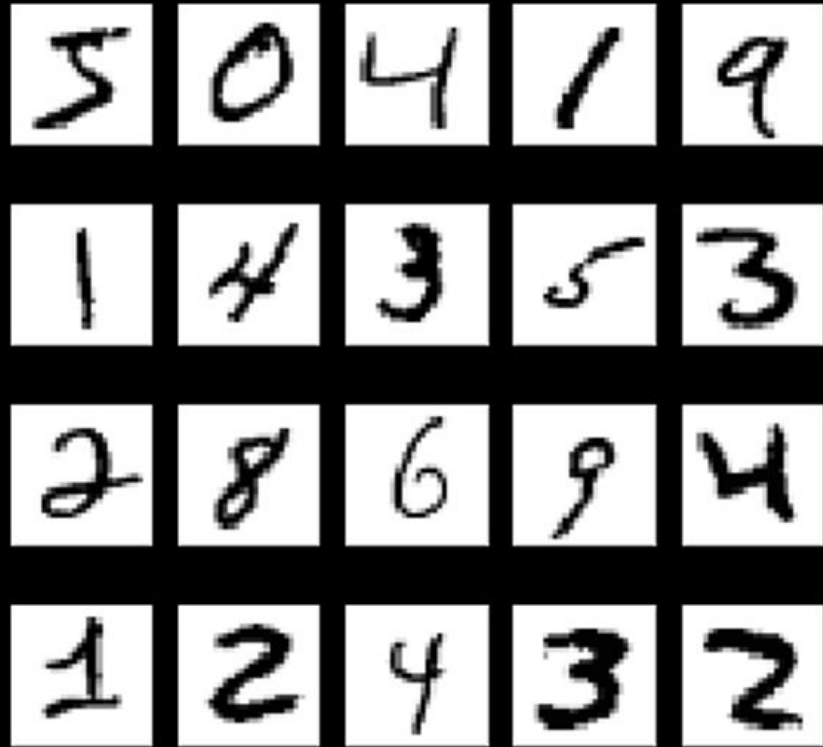
# LOCALITY-SENSITIVE HASH (LSH)

Hash: 101...

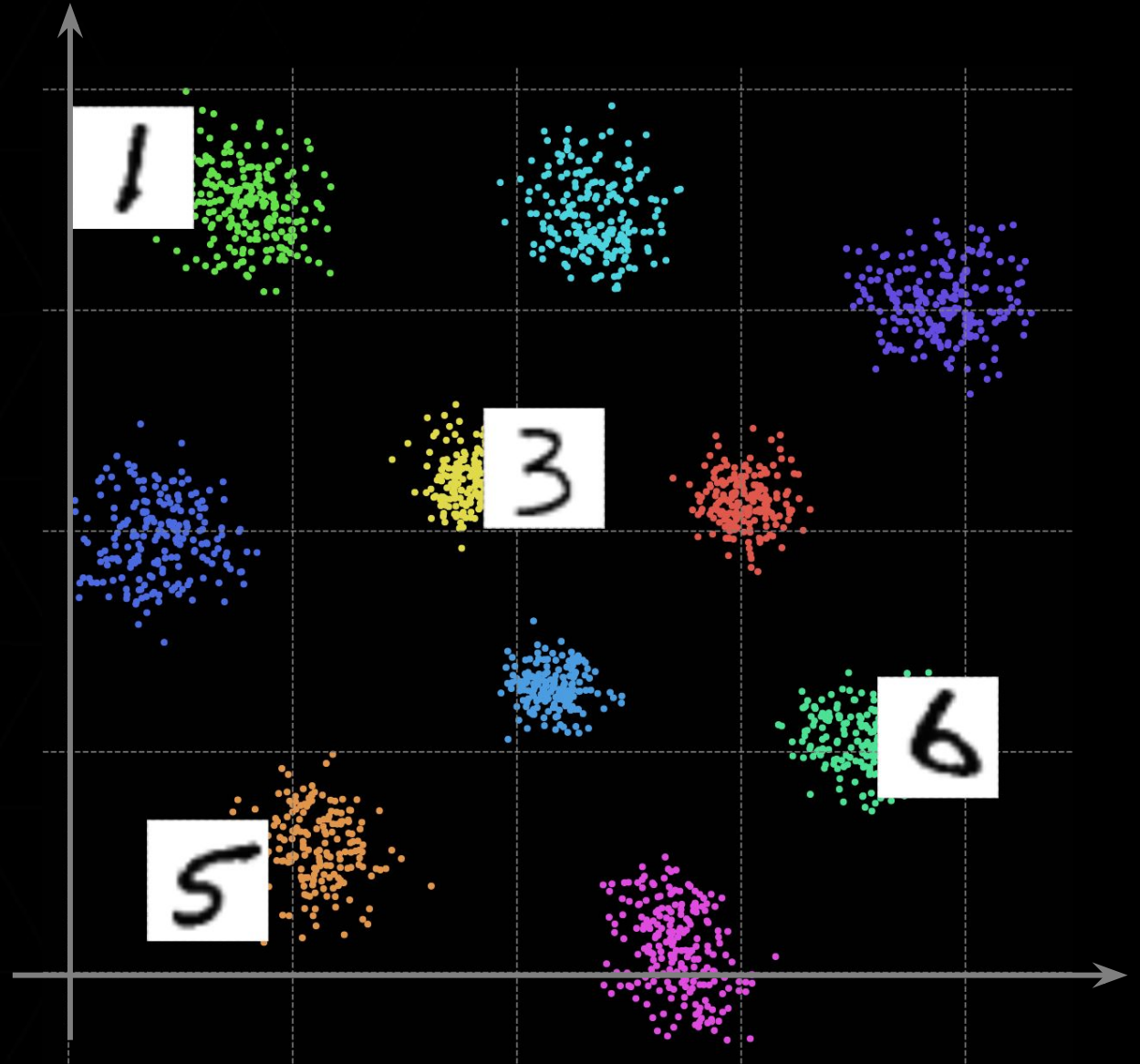
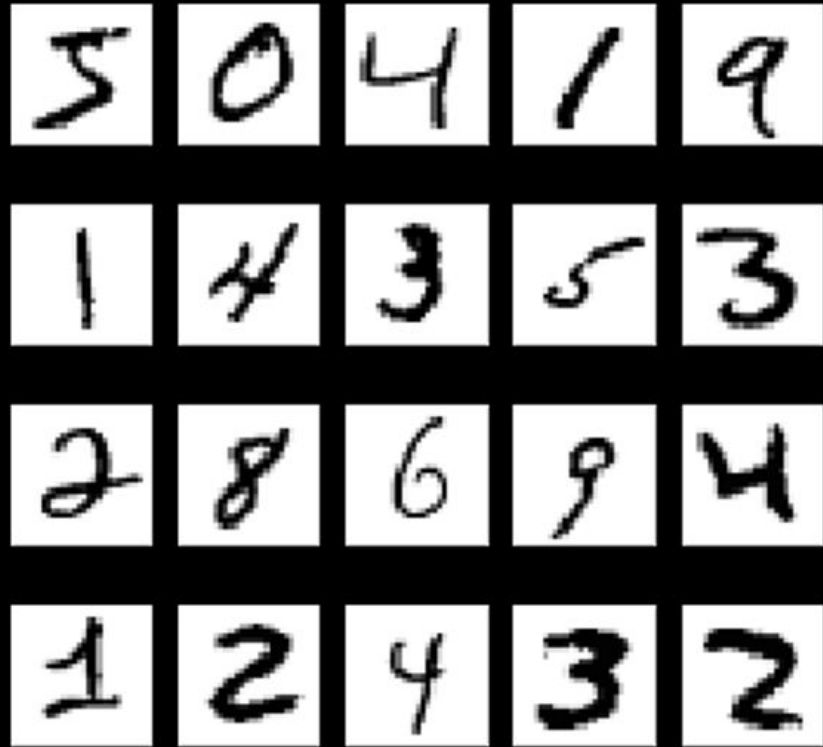




# КЛАСТЕРИЗАЦИЯ ИЗОБРАЖЕНИЙ

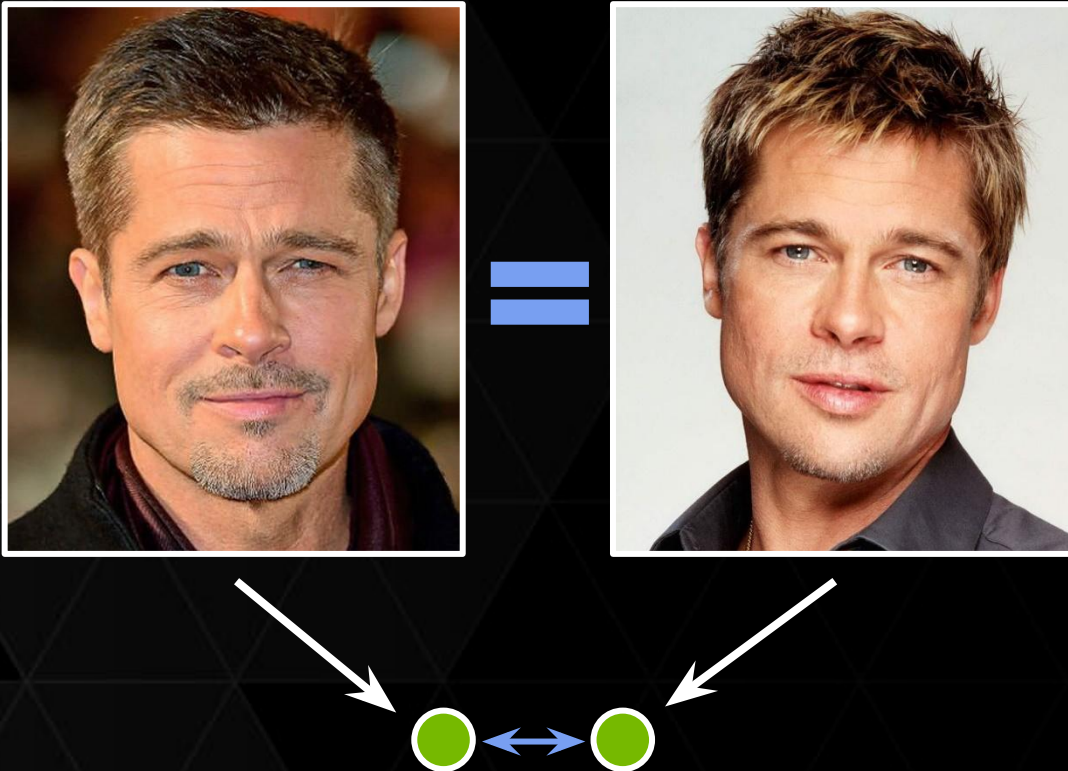


# КЛАСТЕРИЗАЦИЯ ИЗОБРАЖЕНИЙ



# ВЕРИФИКАЦИЯ ЛИЦ

Расстояние меньше порога



# ВЕРИФИКАЦИЯ ЛИЦ

Расстояние меньше порога



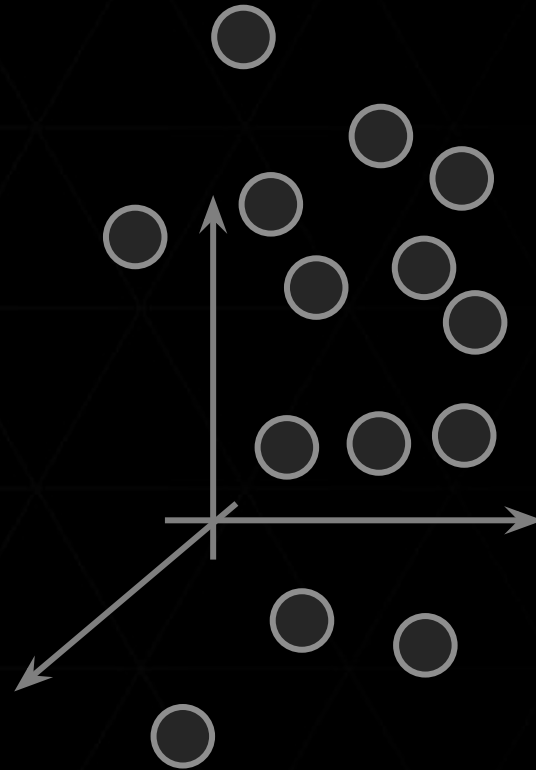
Расстояние больше порога





# ПОИСК ПОХОЖИХ ИЗОБРАЖЕНИЙ

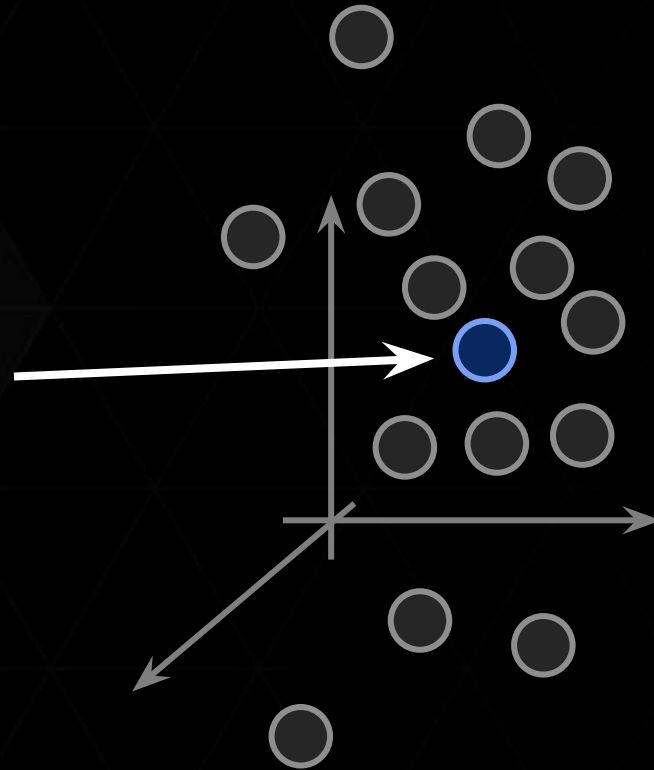
Запрос





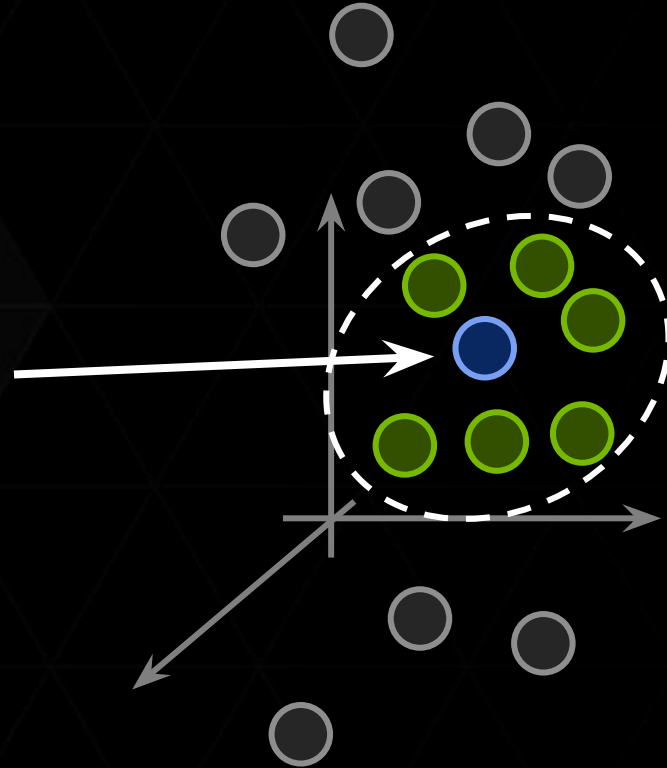
# ПОИСК ПОХОЖИХ ИЗОБРАЖЕНИЙ

Запрос



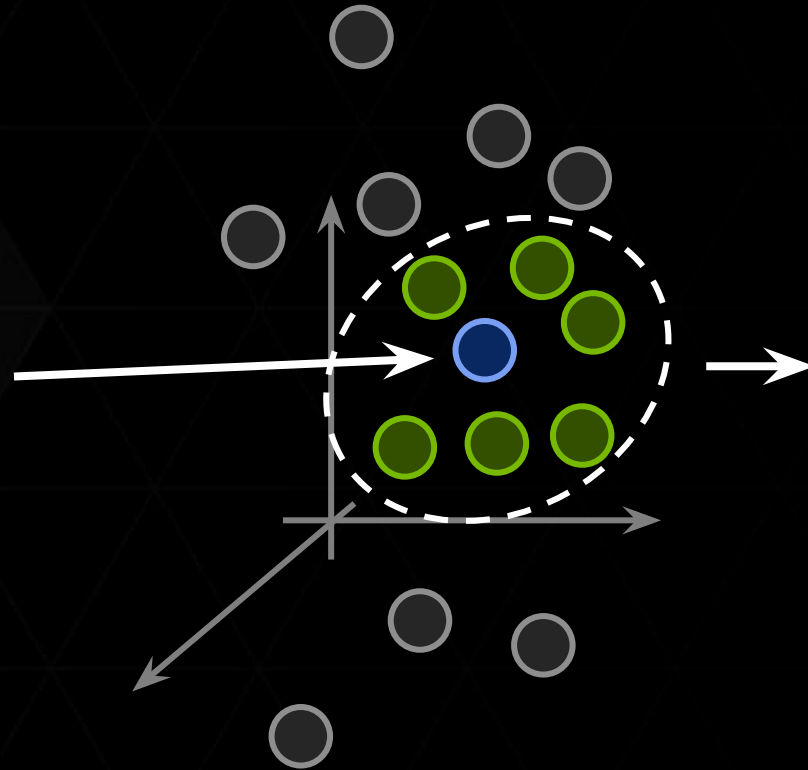
# ПОИСК ПОХОЖИХ ИЗОБРАЖЕНИЙ

Запрос



# ПОИСК ПОХОЖИХ ИЗОБРАЖЕНИЙ

Запрос

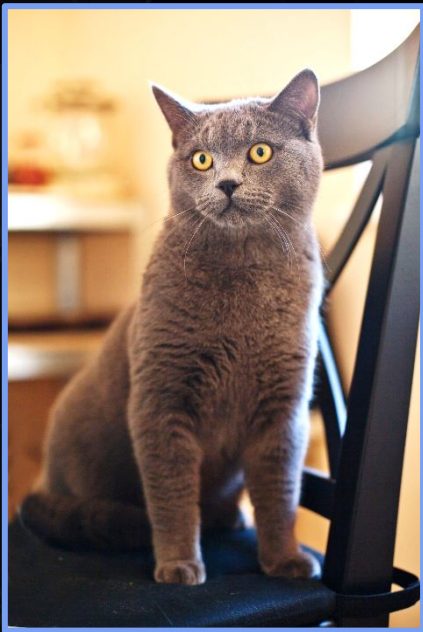


Результат

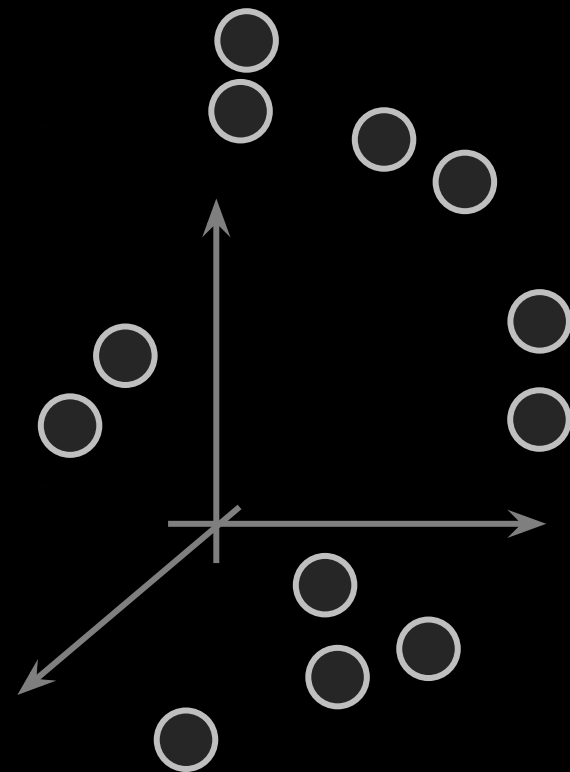




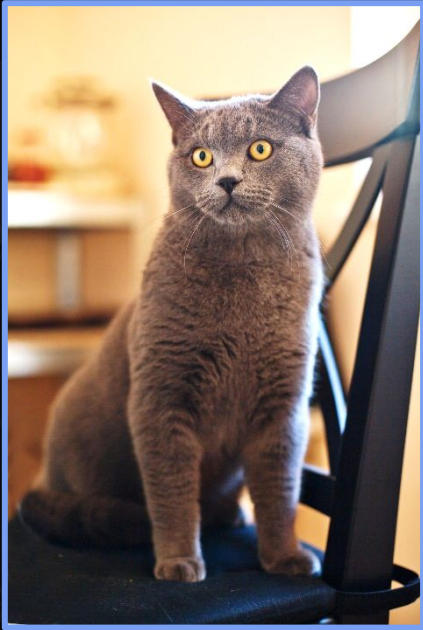
# МУЛЬТИМОДАЛЬНЫЕ ЭМБЕДДИНГИ



«Кот сидит  
на стуле»

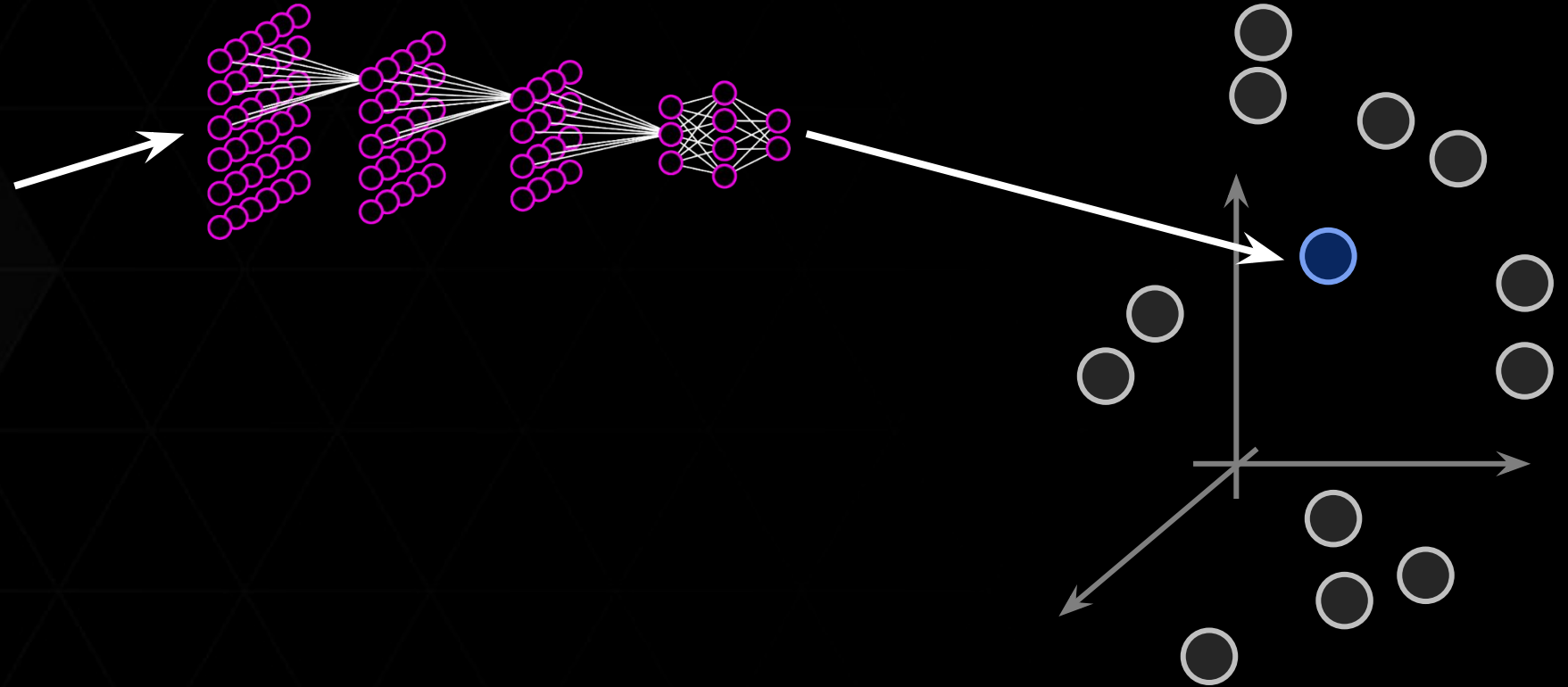


# МУЛЬТИМОДАЛЬНЫЕ ЭМБЕДДИНГИ



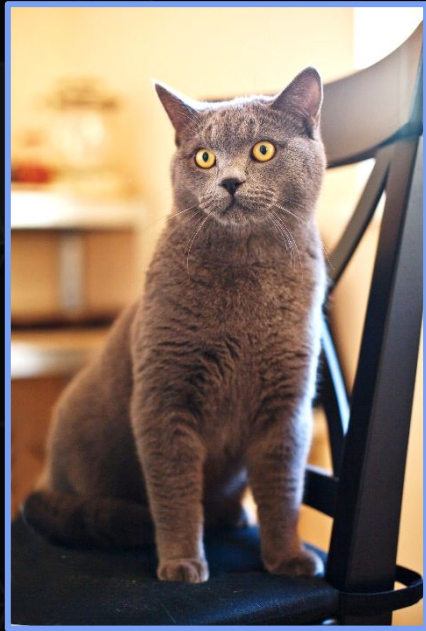
«Кот сидит  
на стуле»

Свёрточная нейросеть



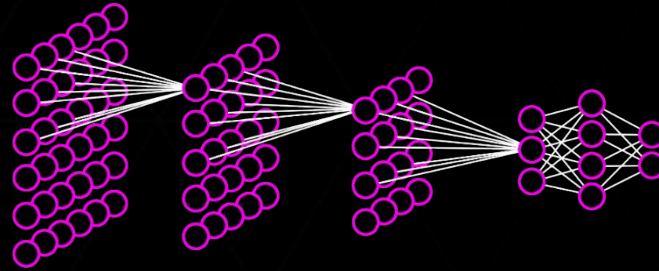


# МУЛЬТИМОДАЛЬНЫЕ ЭМБЕДДИНГИ

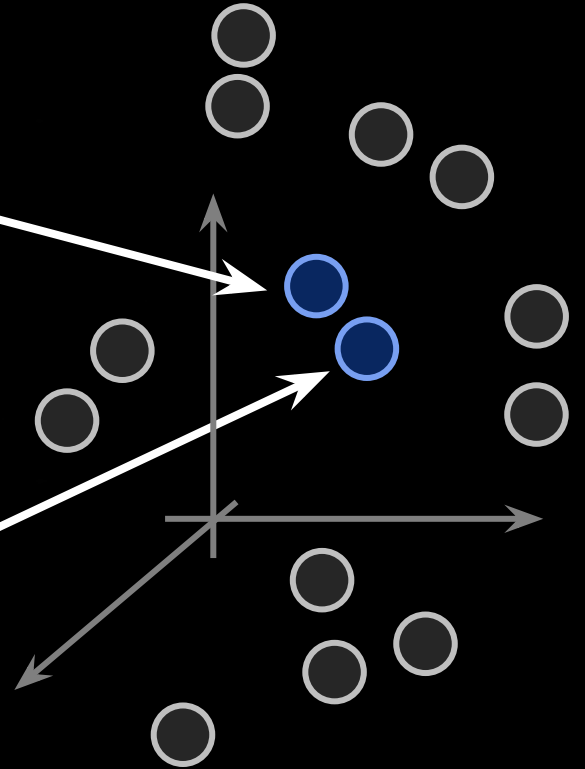
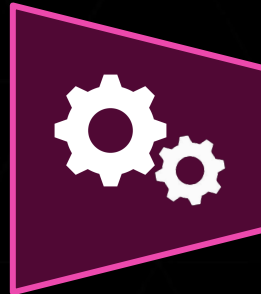


«Кот сидит  
на стуле»

Свёрточная нейросеть



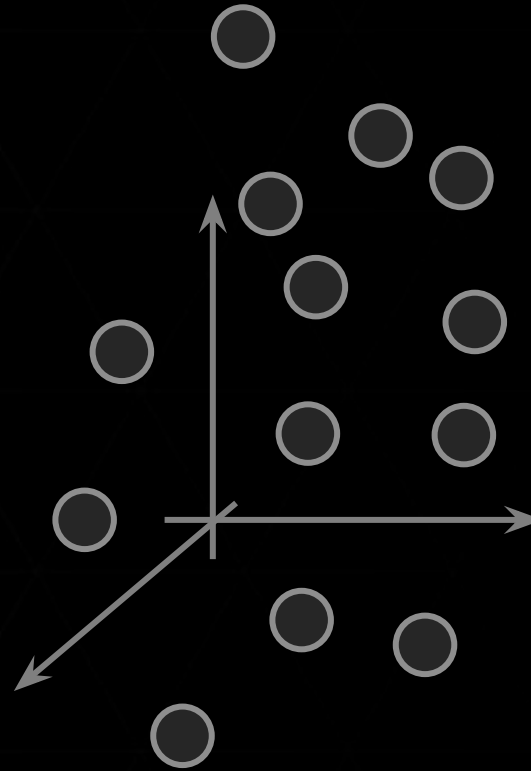
Рекуррентная  
нейросеть



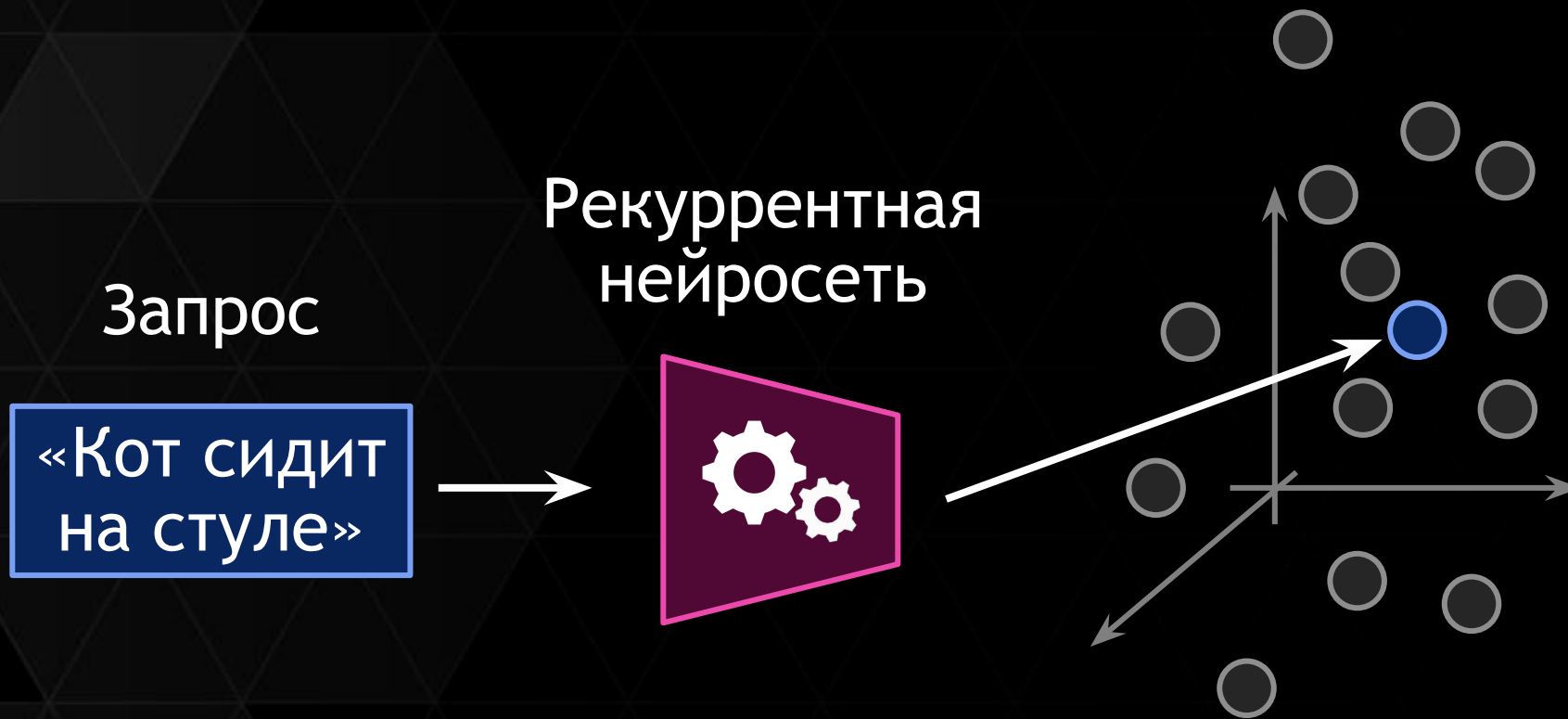
# СЕМАНТИЧЕСКИЙ ПОИСК ИЗОБРАЖЕНИЙ

Запрос

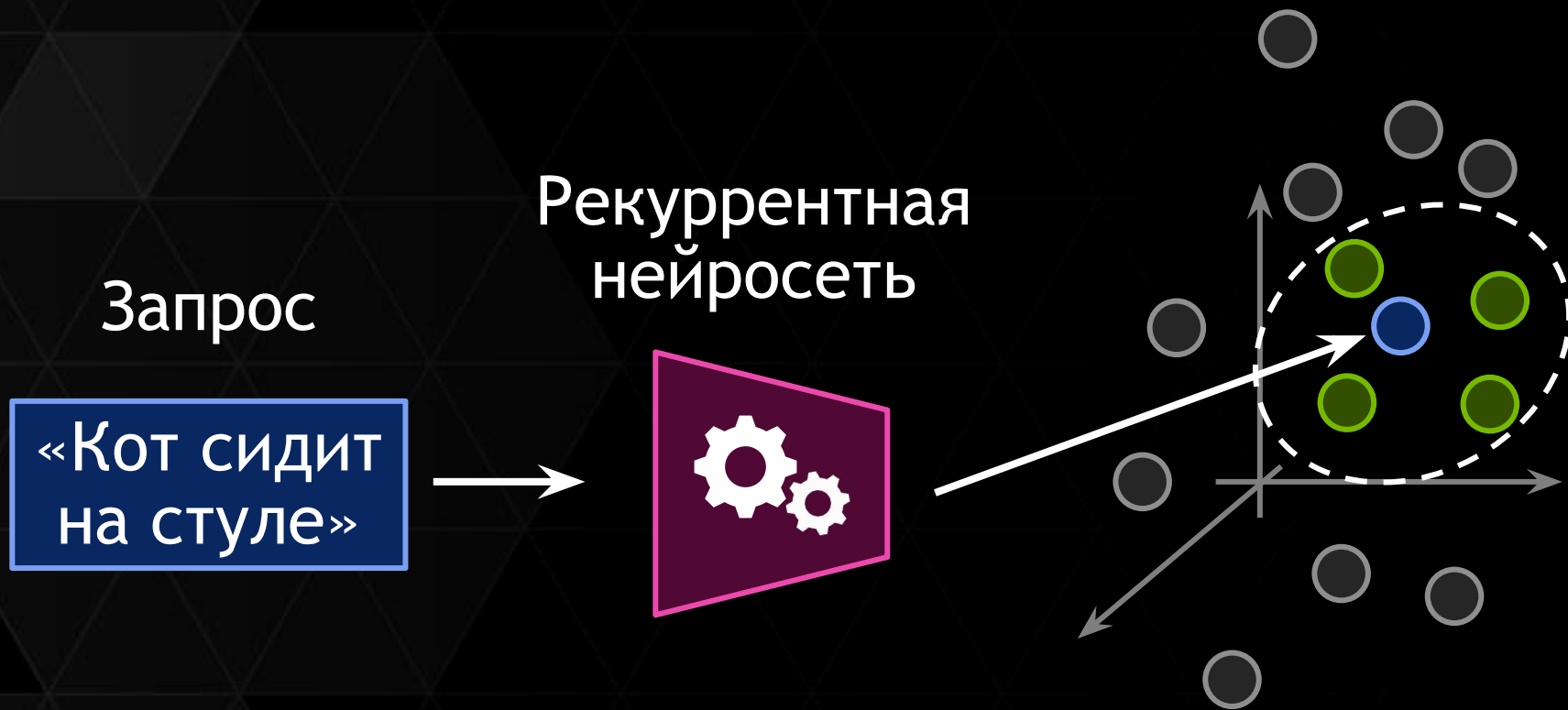
«Кот сидит  
на стуле»



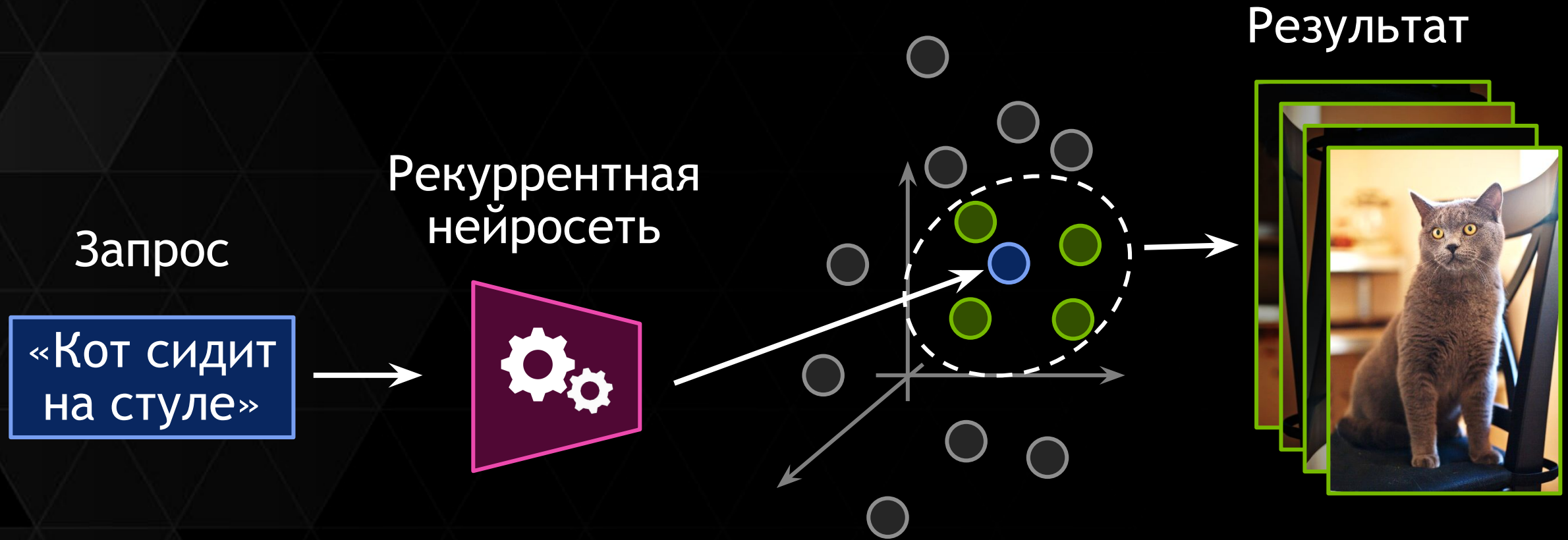
# СЕМАНТИЧЕСКИЙ ПОИСК ИЗОБРАЖЕНИЙ



# СЕМАНТИЧЕСКИЙ ПОИСК ИЗОБРАЖЕНИЙ



# СЕМАНТИЧЕСКИЙ ПОИСК ИЗОБРАЖЕНИЙ



The background is a dark blue field with a network of thin, light green lines connecting various points. Some points are small, bright green dots, while others are larger, fainter blue circles. The lines crisscross the frame, creating a sense of interconnectedness and movement.

СПАСИБО