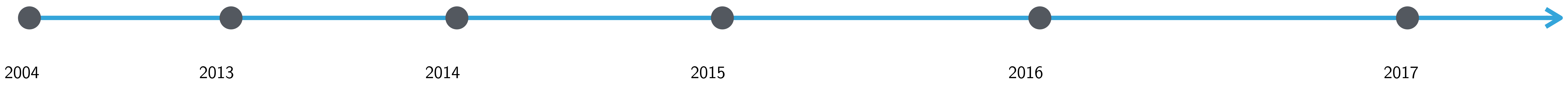
ADVERSARIAL MACHINE LEARNING

10

TAXONOMY OF ADVERSARIAL ATTACKS

- Based on the knowledge of the adversary
  - White Box Attack
    - Adversary has access to the trained model and weights
  - Black Box Attack
    - Adversary has only access to the output of the model e.g. API

TIMELINE



TAXONOMY OF ADVERSARIAL ATTACKS

- Based on the knowledge of the adversary
  - White Box Attack
    - Adversary has access to the trained model and weights
  - Black Box Attack
    - Adversary has only access to the output of the model e.g. API