

Hoofdstuk 1

Classificatie

1.1 Wat is classificatie?

Classificatie in machine learning is een vorm van supervised learning waarbij het doel is om een input toe te wijzen aan een van de vooraf gedefinieerde klassen. De klassen worden vaak ook aangeduid als *label* of *categorie*. Het ultieme doel van classificatie is om te bepalen in welke categorie nieuwe gegevens zullen vallen.

We passen dit even toe op onze tumordataset. We willen aan de hand van classificatie onderzoeken of wat de aard van een tumor is. Er zijn dus twee mogelijke categorieën waartoe een tumor behoort: die van de goedaardige tumoren of die van de kwaadaardige. Omdat er slechts twee mogelijke klassen zijn, is er sprake van *binaire classificatie*. Er bestaan ook echter ook classificatietechnieken waar meerdere klassen aan bod komen.

1.2 Enkele toepassingen

- **Beeldherkenning:** Classificatie wordt vaak gebruikt in beeldherkenningstoepassingen, zoals het identificeren van objecten in foto's of video's.
- **Tekstclassificatie:** Hier wordt classificatie gebruikt om tekst te categoriseren, bijvoorbeeld het identificeren van spam-e-mails of het toewijzen van artikelen aan specifieke onderwerpen.
- **Medische diagnose:** Naast het eerder genoemde voorbeeld van tumorclassificatie, wordt classificatie ook toegepast op het diagnosticeren van andere medische aandoeningen op basis van verzamelde gegevens.