

১. ডেটা প্রিপারেশন: প্রথমে ডেটাকে পরিষ্কার এবং প্রস্তুত করা হয়। এর মধ্যে ডেটা ক্লিনিং, নরমালাইজেশন, এবং ডেটা অগমেন্টেশনের কাজ থাকে, যাতে মডেল আরও ভৈচিত্র্যময় ডেটার সাথে ট্রেনিং পায়।

CNN মডেলে আউগমেন্টেশন ব্যবহার করা হয় মূলত ডেটার বৈচিত্র্য বাড়ানোর জন্য। এইভাবে মডেল বিভিন্ন ধরনের ডেটা দেখতে পারে, যা এর জেনারলাইজেশন ক্ষমতা বাড়ায়। এর ফলে মডেল নতুন নতুন ওয়্যারিয়েশনের সাথে খাপ খাইয়ে নিতে শিখে, এবং রিয়েল-ওয়ার্ল্ড ডেটার সাথে ভালভাবে অ্যাডাপ্ট করতে পারে। এতে মডেলের পারফরমেন্স অনেকটাই উন্নত হয়।

জেনারেলাইজেশন বলতে আসলে বোঝায় মডেলটি নতুন, আগে দেখা না থাকা ডেটার উপর কেমন পারফরমেন্স করছে। অর্থাৎ, মডেলটি যে ট্রেনিং ডেটা থেকে শিখেছে, সেটা ঠিক একইভাবে নতুন ডেটার জন্যও প্রয়োগ করতে পারছে কিনা। যদি সেটা ভালোভাবে করতে পারে, তাহলে মডেলটি ভালোভাবে জেনারেলাইজড বলা হয়।

২. মডেল আর্কিটেকচার ডিজাইন: কনভলিউশনাল লেয়ার, পুলিং লেয়ার, ফ্ল্যাটেন লেয়ার ইত্যাদি দিয়ে মডেলটি ডিজাইন করা হয়।

৩. হাইপারপ্যারামিটার টিউনিং: অটো-টিউনিংয়ের মাধ্যমে সেরা হাইপারপ্যারামিটার সেট নির্ধারণ করা হয়।

অটো-টিউনিং হলো এমন এক প্রক্রিয়া, যেখানে মডেলের হাইপারপ্যারামিটারগুলো অটোমেটিকভাবে টিউন করা হয়, যাতে মডেলটি তার সেরা পারফরমেন্স দিতে পারে। এই প্রক্রিয়ায় বিভিন্ন হাইপারপ্যারামিটার সেটিংস পরীক্ষা করে, আর সেগুলো এমনভাবে নির্বাচন করে যে মডেলের পারফরমেন্স সর্বোচ্চ হয়। এর ফলে মডেল টিউনিংয়ের জন্য আলাদা করে সময় বা চেষ্টা দিতে হয় না, এটা অনেকটাই স্বয়ংক্রিয় হয়ে যায়।

অটো-টিউনিং হাইপারপ্যারামিটার সেট করে দেয়, যাতে মডেলটি নিজে সেরা পারফরমেন্স দিতে পারে

অটো-টিউনিং এবং অগমেন্টেশন দুটোই মডেলের পারফরমেন্স বাড়াতে কাজ করে, তবে এদের কাজের ধরন আলাদা। অটো-টিউনিং হাইপারপ্যারামিটার সেট করে দেয়, যাতে মডেলটি নিজে সেরা পারফরমেন্স দিতে পারে, আর অগমেন্টেশন ডেটার বৈচিত্র্য বাড়ায়। এই দুইটি প্রক্রিয়া একসাথে ব্যবহার করলে মডেল আরও ভালোভাবে ট্রেনিং পায় এবং তার পারফরমেন্স অনেকটাই উন্নত হয়।

৪. মডেল কম্পাইল এবং ট্রেনিং: হাইপারপ্যারামিটার সেট করার পর মডেলটি কম্পাইল করে ট্রেন করা হয় ট্রেনিং ডেটা দিয়ে।

মডেল কম্পাইল করা হয় মডেলটিকে ট্রেনিংয়ের জন্য প্রস্তুত করার জন্য, যাতে সেটি লস ফাংশন, অপটিমাইজার ইত্যাদি জানে।

কলব্যাক সেটআপ: ট্রেনিংয়ের সময় আর্লি স্টপিং, মডেল সেভ ইত্যাদি কলব্যাকস সেটআপ করা হয়, যাতে ট্রেনিং আরও কার্যকরী হয়। কলব্যাকস সেটআপ করা হয় ট্রেনিংয়ের সময় বিশেষ কিছু কাজ যেমন ইয়ারলি স্টপিং, মডেল সেভ করা ইত্যাদি অটোমেটিকভাবে করার জন্য।

৫. ভ্যালিডেশন এবং ইভালুয়েশন: ট্রেনিংয়ের সময় ভ্যালিডেশন ডেটা দিয়ে মডেলের পারফরমেন্স মনিটর করা হয়, এবং যদি পারফরমেন্স স্থির হয়ে যায়, তাহলে আর্লি স্টপিং ব্যবহার করা হয়। ট্রেনিং কার্ড প্লট করলে ট্রেনিং এবং ভ্যালিডেশন লস এবং অ্যাকুরেস দেখা যায়, যা থেকে বোঝা যায় মডেল কেমন পারফর্ম করছে। আর ইনফারেন্স হলো ট্রেনিং শেষ হওয়ার পর মডেলটি নতুন ডেটার উপর প্রেডিকশন করা।

৬. ইনফারেন্স: ট্রেনিং এবং সেভ করার পর নতুন ডেটার উপর মডেলটি প্রেডিকশন করতে প্রস্তুত হয়। সহজ কথায়, মডেলটা যখন আগে থেকে শিখে গিয়েছে, তখন সেটা নতুন ইনপুট দিয়ে তার থেকে আউটপুট বা প্রেডিকশন দেয়। ইনফারেন্স সাধারণত ট্রেনিংয়ের তুলনায় অনেক দ্রুত হয় এবং মডেলকে বাস্তব জীবনের ডেটার সাথে কাজ করতে সাহায্য করে।