**Review Learning Image Classification Model Deployment**

1. Reinforcement Learning
   1. Pendahuluan Reinforcement Learning

Jika Anda belum tahu, bot pada sebuah game adalah program yang memainkan game itu sendiri tanpa adanya campur tangan manusia manusia. AlphaGo dilatih menggunakan sebuah teknik yaitu reinforcement learning.

* 1. Pengenalan Reinforcement Learning

Reinforcement Learning (RL) adalah salah satu cabang ML yang paling menarik saat ini dan juga salah satu cabang ML paling tua. RL telah ada sejak tahun 1950-an dan dipakai dalam permainan seperti TD-Gammon. Namun saat itu RL tidak terlalu menarik perhatian publik.

Bagaimana cara kerja RL? Pada Reinforcement Learning sebuah perangkat lunak yang disebut agen membuat pengamatan terhadap lingkungan dan melakukan aksi-aksi yang sesuai dari pengamatan yang dilakukan.

Tujuan dari agen adalah untuk melakukan tindakan-tindakan yang terus menambah reward/hadiah.

Analoginya seperti teorema hadiah dan hukuman. Kita sebagai manusia cenderung melakukan hal yang membuat kita senang (hadiah) dan menghindari kegiatan yang membuat kita menderita (hukuman). Demikian halnya dengan agen pada RL yang berusaha untuk mendapatkan sebanyak-banyaknya hadiah dan meminimalisir hukuman.

* 1. Policy Search Algorithm

Algoritma yang digunakan agen untuk menentukan aksi apa yang harus dilakukan disebut dengan policy/kebijakan. Sebuah policy dapat berupa sebuah model ML yang menerima masukan berupa pengamatan terhadap lingkungannya. Lantas menghasilkan output berupa aksi yang harus ia kerjakan.

Selain itu, ada salah satu teknik yang dapat diaplikasikan pada policy search yaitu genetic algorithm. Algoritma genetik sangat berguna terutama dalam pencarian policy terbaik dengan waktu yang lebih cepat.

* 1. Deep Reinforcement Learning

Jaringan saraf sangat cocok dipakai pada situasi di mana aksi atau policy yang dilakukan banyak sekali. Pada awal sekali, weight dari jaringan saraf akan diambil secara acak. Melalui hadiah dan penalti yang didapat dari setiap aksinya, jaringan saraf menyesuaikan weight-nya. Hal ini sangat mirip dengan backpropagation pada supervised learning. Bedanya, pada supervised learning telah terdapat target yang ingin dicapai misalnya gambar yang telah dilabeli. Sedangkan RL sangat bergantung feedback yang didapat dari lingkungan untuk menentukan aksi yang harus dikerjakan.

* 1. TF-Agents dan Deep Q-Network

TF-agent adalah sebuah library Reinforcement Learning dari Tensorflow yang dikembangkan google dan dirilis ke publik pada tahun 2018. TF-agents menyediakan lingkungan pengembangan RL seperti yang disediakan oleh OpenAI Gym, serta mendukung simulasi fisika 3D dari library PyBullet, dan mendukung simulasi pada lingkungan 3D yang disediakan pada library unity ML agents.

1. Deployment
   1. Intro

Bagaimana cara men-deploy model yang sudah kita buat pada Web Browser atau mobile (iOS dan Android). Men-deploy model bukan hanya memasangkan model pada Web Browser atau mobile saja, melainkan juga memonitor/mengawasi, dan memperbarui model tersebut.

* 1. Pengenalan Deployment

Alur proses produksi dimulai dari proses deployment. Setelah kita memiliki model yang sudah dilatih terhadap suatu dataset, model tersebut kita deploy sesuai platform yang ingin kita gunakan. Proses deployment pada machine learning adalah suatu cara di mana Anda menggunakan model yang telah Anda latih (trained model) dan membuatnya menjadi suatu entitas yang dapat memberikan hasil prediksi dari live data. Berikut adalah alur deployment sebuah project machine learning.

Model machine learning di-deploy dengan menggunakan 3 cara utama :

* Menggunakan model server
* Browser
* Edge devices
  1. Pengenalan TensorFlow.js

Tensorflow.js adalah sebuah framework yang kompatibel dengan TensorFlow API. TensorFlow.js menggunakan model yang telah dibuat dengan merubah format model menjadi JSON file. Pada level atas TensorFlow.js API, kita akan dihadapkan dengan Layers API. Jika kita familiar dengan penggunaan layer-layer pada Keras, kita dapat dengan mudah menggunakan Layers API. Di bawah layers API, ada Core API yang menangani model dari TensorFlow. Core API juga mengimplementasikan operasi graf pada level lebih kompleks seperti deklarasi tensor (data input), operasi pada tensor, memori, eksekusi fungsi, dan lain-lain.

* 1. Deploy Model NLP pada Browser

Menjalankan model yang telah kita buat untuk melakukan tugas prediksi pada lingkup pemrosesan bahasa alami (Natural Language Processing) yang dijalankan pada Web Browser menggunakan TensorFlow.js. Adapun pengetahuan tambahan tentang HTML, CSS, JavaScript, atau Chrome Dev Tools dapat membantu anda memahami sub-modul ini dengan baik.

* 1. Deploy Model ML ke Android dan iOS dengan TF-Lite

TensorFlow lite adalah seperangkat alat yang membantu programmer untuk menjalankan TensorFlow pada mobile, dan perangkat IoT. Machine Learning pada mobile dan perangkat IoT membantu manusia dalam menjalankan kehidupan sehari-hari agar lebih mudah

Ada beberapa keuntungan jika kita men-deploy Machine Learning pada mobile:

* Penggunaan Machine Learning bisa tidak memakai koneksi internet, sehingga dapat digunakan dimanapun dan menjaga privasi pengguna.
* Latency dari perangkat mobile jauh lebih kecil daripada komputer.

TensorFlow Lite terdiri atas 2 komponen utama:

* TensorFlow Lite Interpreter → Menjalankan model TensorFlow pada banyak hardware yang berbeda, termasuk mobile dan microcontroller.
* TensorFlow Lite Converter → Mengonversikan model TensorFlow menjadi bentuk yang efisien agar dapat digunakan oleh Interpreter, juga mengoptimasi performance dari model tersebut.
  1. Federated Learning

Federated learning merupakan konsep sistem pembelajaran mesin modern yang mengutamakan privasi serta kecerdasan sistem. Sampai saat tulisan ini dibuat, federated learning masih berada dalam area riset aktif yang masih dapat berkembang namun sudah berhasil untuk mengatasi permasalahan di atas. Federated learning memungkinkan pengembang bereksperimen dengan data-data yang tidak terpusat (decentralized) atau dengan kata lain terdistribusi pada masing-masing client. Data yang terdistribusi ini menarik karena data ini bersifat pribadi dan ada di mana saja. Milyaran perangkat pribadi seperti telepon genggam dan peranti IoT senantiasa menghasilkan data-data baru yang bisa dilatih untuk menciptakan kemudahan bagi para penggunanya.

* 1. Data Pipelines dengan TensorFlow Data Services

TensorFlow Data Services (TFDS) membantu mengatur data pada machine learning menjadi lebih mudah. TFDS mengubah banyak baris kode tersebut menjadi sangat sedikit. Format TFDS juga dapat menyelesaikan masalah-masalah aturan dataset sebelumnya dengan menggunakan platform yang dapat diakses dengan API yang konsisten. Tujuannya adalah mempermudah dan meminimalisir perbedaan pada saat mengakses data.

* 1. Export Data ke Training Pipelines

Mengubah sebuah dataset (berupa CSV atau Pandas DataFrame) menjadi format TFDS dengan cara yang efisien.

* 1. Mempublikasi Model Anda ke TF-Hub

TensorHub (TensorFlow Hub) adalah tempat mencari, mempublikasikan, dan menggunakan reusable machine learning model yang sudah ada. Model yang ada di TensorHub berisi weight dan aset-aset yang dapat digunakan untuk pekerjaan lainnya yang dikenal sebagai transfer learning.

* 1. Pengenalan TensorBoard

TensorBoard dapat diakses baik pada sistem lokal maupun suatu lingkungan hosting. Tensorboard.dev memungkinkan kita membagikan eksperimen kita pada website TensorBoard. Gunanya untuk memudahkan kolaborasi pekerjaan tim, troubleshooting, publikasi, dan lain-lain.

* 1. Memonitor Model

Model monitoring atau pemantauan model pada ML adalah suatu proses pengamatan kinerja model pada fase produksi dilihat dari perspektif operasional dan teknis. Sistem ML tidaklah sama pada sistem biasa pada umumnya. Ini disebabkan dependensi sistem ML bukan hanya pada kode, melainkan juga pada perilaku model yang bergantung sepenuhnya pada data yang dipelajari. Tanpa memahami perilaku model dan data yang dipelajari, kita tidak akan mengerti operasional sistem yang dibuat. Oleh karena itu tanggung jawab pemantauan model pada sistem ML tidak dapat secara otomatis diserahkan ke tim DevOps. Validasi penjaminan operasional yang tepat dari sistem ML ini tidak terlepas dari keterlibatan langsung antara tim DevOps dan data scientist.

* 1. Memperbarui Model yang Telah Di-deploy

Beberapa alasan mengapa kita harus melakukan pembaruan pada model. Seperti berubahnya trend dan selera user, atau kinerja model yang semakin lama semakin buruk. Untuk mengantisipasi hal tersebut amati statistik tentang data pada suatu waktu tertentu guna memahami tren atau pola user. Sedangkan untuk mengamati kinerja model, pengembang dapat mengukur kinerja model terhadap suatu metrik secara berkala baik offline maupun online metrik beserta korelasi di antara keduanya.