

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Setelah melakukan penelitian dalam melakukan *clustering* menggunakan algoritma *hierarchical clustering* pada *DNA barcoding*. Maka dapat diambil kesimpulan sebagai berikut dari penelitian yang telah dilakukan.

1. Algoritma *hierarchical clustering* dapat digunakan untuk melakukan *clustering* dalam kasus *DNA Barcoding* dengan menjalankan beberapa tahap pada model komputasi, yaitu: (i) *data collection*, (ii) *data preprocessing*, (iii) *finding the best distance method*, and (iv) *determining disputed family*.
2. *Pearson correlation* merupakan pendekatan yang paling baik dalam eksperimen ini dalam hal akurasi 99.04% dalam akurasi (*weighted average*) dan 0.015 detik dalam waktu komputasi (*weighted average*) dibanding metode lainnya berdasarkan tahap validasi menggunakan data dari beberapa famili dan genus yang tidak dipermasalahkan oleh para ahli. Metode *Manhattan*, *Pearson*, dan *Spearman* mendapatkan hasil akurasi terbaik pada 99.04%, namun dalam hal waktu komputasi *Pearson* memiliki waktu komputasi tercepat (0.015 detik) dibandingkan dengan *Manhattan* (0.0157 detik) dan *Spearman* (0.0598).
3. Penelitian ini memvalidasi bahwa teknik *machine learning*, lebih spesifiknya *hierarchical clustering* dapat digunakan untuk memastikan permasalahan dalam taksonomi yang tidak dapat diselesaikan menggunakan ciri ciri morfologi dari tumbuhan. Akurasi yang didapatkan dari model *hierarchical clustering* memiliki akurasi yang tinggi diatas 98% dengan waktu komputasi bervariasi tergantung dari *distance method* yang digunakan dengan rata rata 0.03 detik pada data yang digunakan.
4. Dalam studi kasus *Leguminosae* peneliti menemukan bahwa famili *Fabaceae*, *Mimosaceae*, dan *Caesalpiniaceae* seharusnya masuk kedalam satu famili yang sama. Hal ini didapatkan dari hasil interpretasi yang

Muhammad Iqbal Zain, 2023

IMPLEMENTASI HIERARCHICAL CLUSTERING DALAM DNA BARCODING UNTUK MENENTUKAN TAKSONOMI TUMBUHAN

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

dilakukan oleh ahli dalam bidang Biologi berdasarkan visualisasi *dendrogram* yang dihasilkan.

5.2 Saran

Dalam pelaksanaan penelitian ini, penulis menyadari bahwa masih banyak kekurangan yang dilakukan dalam penelitian ini. Adapun saran yang dapat dilakukan pada penelitian selanjutnya adalah sebagai berikut.

1. *Distance method* yang digunakan dalam penelitian ini hanya menggunakan *mathematical distance measures*, peneliti belum membandingkan dengan *pairwise distance methods* seperti *Kimura 2-parameter (K2P)* dan *Jukes and Cantor*.
2. Penelitian ini juga belum membandingkan hasil yang didapatkan dengan pendekatan biologis seperti *electrophoretic analysis* untuk melakukan *clustering* terhadap data *barcode* DNA dari spesies.
3. Pendekatan *machine learning* yang digunakan mungkin tidak konsisten tergantung pada *library* ketika mendapatkan update atau penyesuaian terhadap parameter yang mereka gunakan. Yang mana permasalahan ini tidak ditemukan jika menggunakan metode tradisional.
4. Terdapat kekurangan juga terhadap hasil interpretasi yang dihasilkan pada visualisasi *dendrogram* yang harus divalidasi terlebih dahulu melalui ahli dalam bidang biologi.
5. Metode *clustering* lainnya dapat digunakan sebagai pengganti *hierarchical clustering* yang digunakan dalam penelitian ini, seperti *K-means*, *DBSCAN*, *Gaussian Mixture*, dll.
6. Penelitian lanjutan juga dapat menggunakan *Deep Learning* sebagai alternatif pendekatan yang lebih *advance* dalam bidang *machine learning*.
7. Data yang digunakan dalam penelitian ini relatif sedikit untuk memvalidasi, sekitar 75 data DNA digunakan untuk sekali eksperimen. Data yang lebih banyak mungkin dapat berpengaruh dalam memvalidasi metode mana yang terbaik.

8. Ahli yang merepresentasikan hasil visualisasi dendrogram pada Gambar 4.19 hanya memaparkan hasil interpretasinya saja, belum menjelaskan bagaimana batasan dan standard yang ada dari interpretasinya dalam bidang Biologi. Hal ini dapat ditambahkan pada penelitian serupa yang menggunakan *Hierarchical clustering*.