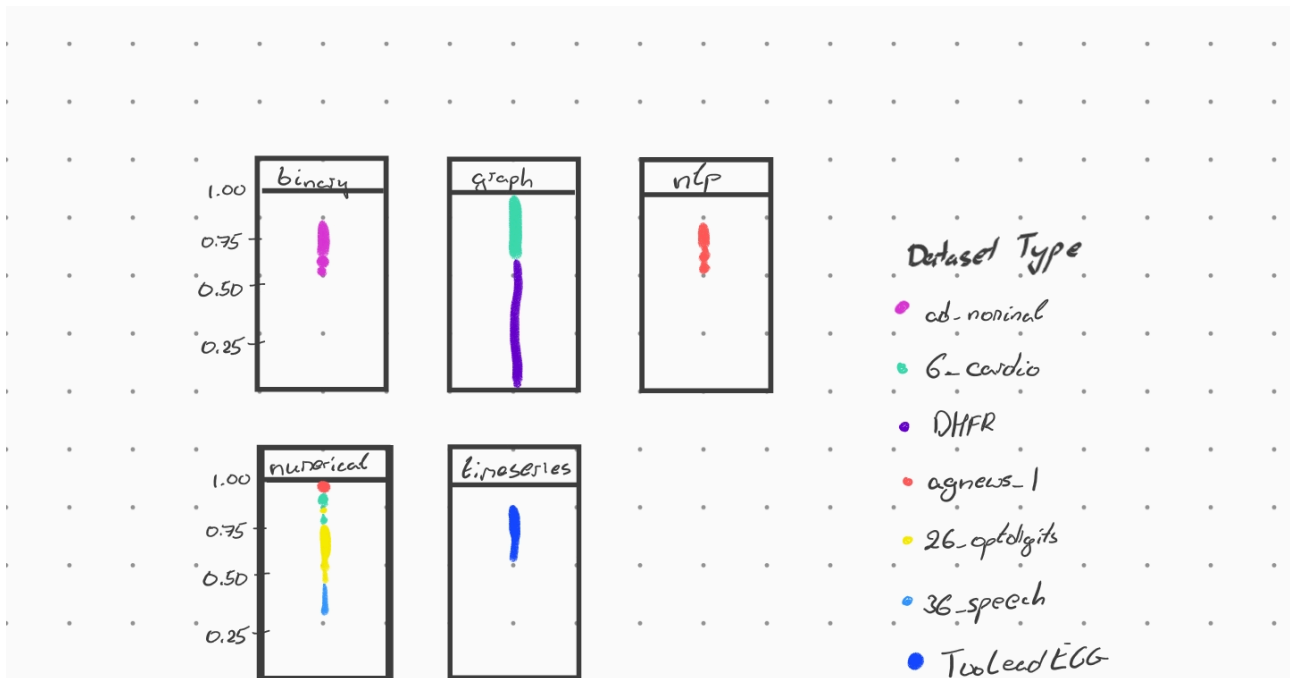


Report 2

Prepared by:

- Maria Musiał - 156062
- Jakub Liszyński - 156060

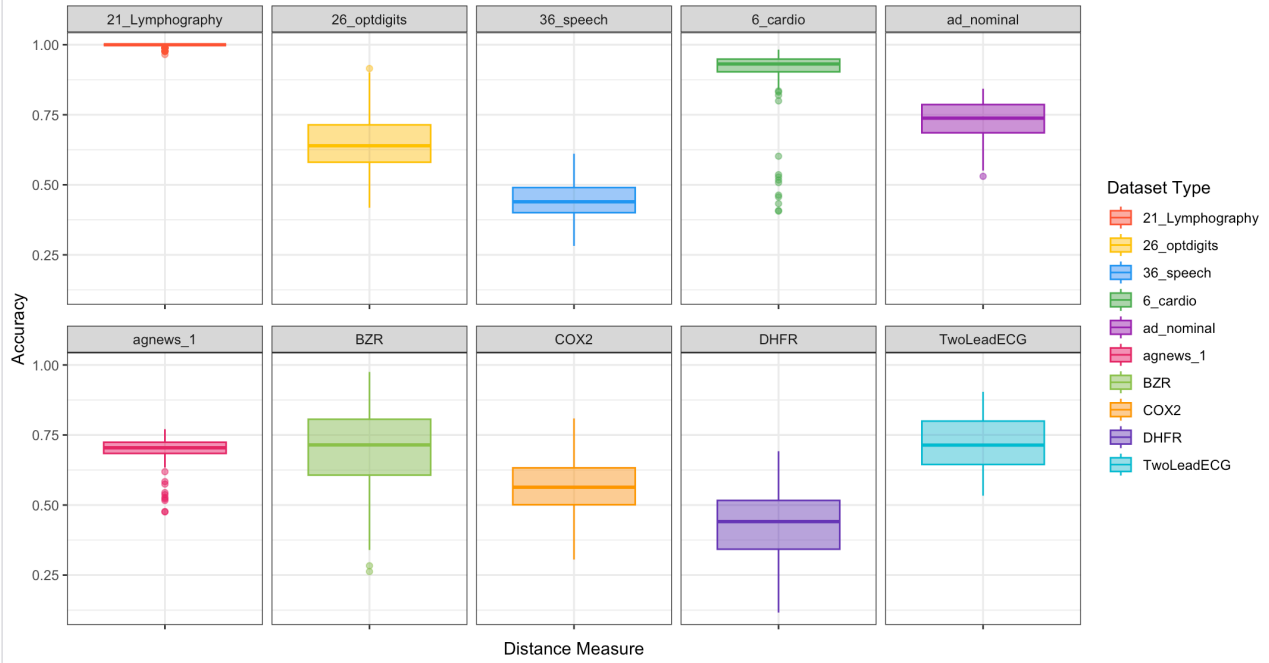
Sketch of one of the graphs



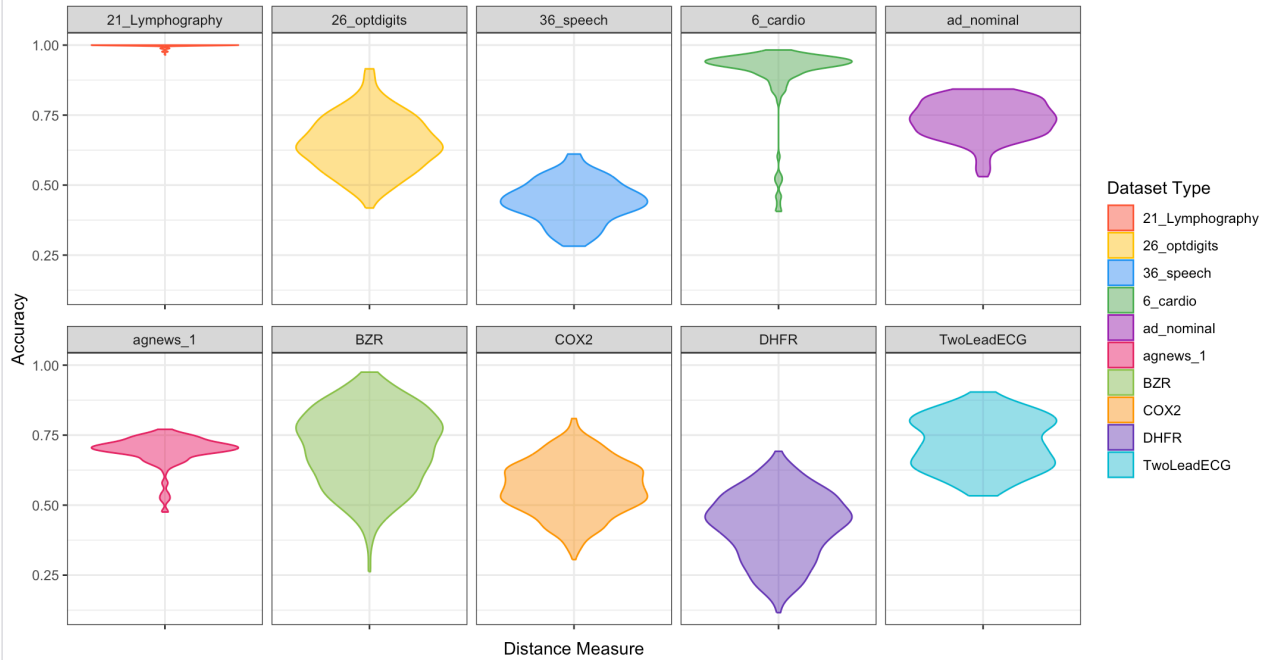
Preliminary plots



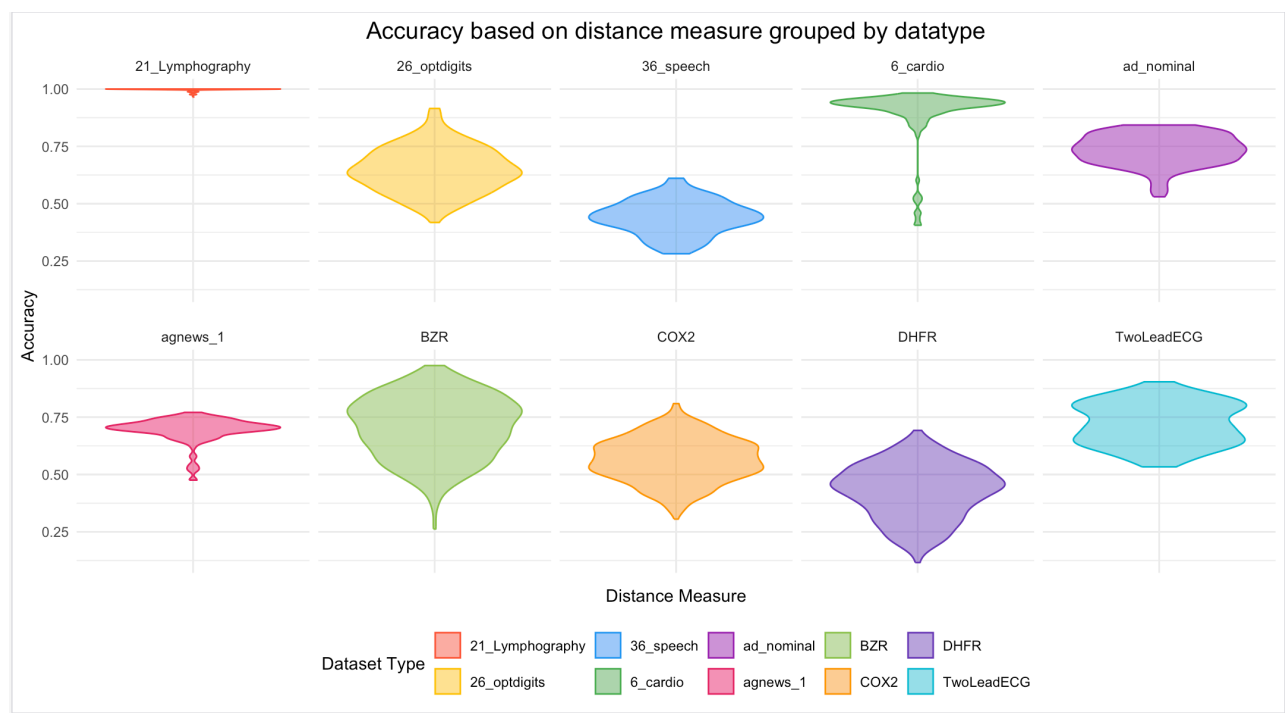
Accuracy based on distance measure grouped by datatype



Accuracy based on distance measure grouped by datatype



Final plot



Code in R

```

12 colour_palette <- c("#FF5733", "#FFC300", "#2196F3", "#4CAF50", "#9C27B0", "#E91E63", "#8BC34A", "#FF9800", "#673AB7", "#9E9E9E")
13 # point
14 ggplot(dfmain, aes(x = distances, y = auc, color = factor(dataset_name), fill = factor(dataset_name))) +
15   geom_point(alpha=0.5) + theme_bw() +
16   scale_color_manual(values = colour_palette, name="Dataset Type") +
17   scale_fill_manual(values = colour_palette, name="Dataset Type") +
18   labs(title = "Accuracy based on distance measure grouped by datatype", x = "Distance Measure", y = "Accuracy") +
19   facet_wrap(~ factor(dataset_name), nrow = 2, scales="free_x") +
20   theme(axis.text.x = element_text(angle = 70, hjust = 1), plot.title = element_text(size=15, hjust=0.5))
21 # boxplot
22 ggplot(dfmain, aes(x = distances, y = auc, color = factor(dataset_name), fill = factor(dataset_name))) +
23   geom_boxplot(alpha=0.5) + theme_bw() +
24   scale_color_manual(values = colour_palette, name="Dataset Type") +
25   scale_fill_manual(values = colour_palette, name="Dataset Type") +
26   labs(title = "Accuracy based on distance measure grouped by datatype", x = "Distance Measure", y = "Accuracy") +
27   facet_wrap(~ factor(dataset_name), nrow = 2, scales="free_x") +
28   theme(axis.text.x = element_text(angle = 70, hjust = 1), plot.title = element_text(size=15, hjust=0.5))
29 # violin
30 ggplot(dfmain, aes(x = distances, y = auc, color = factor(dataset_name), fill = factor(dataset_name))) +
31   geom_violin(alpha=0.5) + theme_bw() +
32   scale_color_manual(values = colour_palette, name="Dataset Type") +
33   scale_fill_manual(values = colour_palette, name="Dataset Type") +
34   labs(title = "Accuracy based on distance measure grouped by datatype", x = "Distance Measure", y = "Accuracy") +
35   facet_wrap(~ factor(dataset_name), nrow = 2, scales="free_x") +
36   theme(axis.text.x = element_text(angle = 70, hjust = 1), plot.title = element_text(size=15, hjust=0.5))
37 # final
38 ggplot(dfmain, aes(x = distances, y = auc, color = factor(dataset_name), fill = factor(dataset_name))) +
39   geom_violin(alpha = 0.5) +
40   theme_minimal() +
41   scale_color_manual(values = colour_palette, name = "Dataset Type") +
42   scale_fill_manual(values = colour_palette, name = "Dataset Type") +
43   labs(title = "Accuracy based on distance measure grouped by datatype", x = "Distance Measure", y = "Accuracy") +
44   facet_wrap(~ factor(dataset_name), nrow = 2, scales = "free_x") +
45   theme(axis.text.x = element_text(angle = 70, hjust = 1), plot.title = element_text(size = 15, hjust = 0.5), legend.position = "bottom")

```