

Relatório para Laboratório 1: “Impactos da Representação”

Nome: Douglas Affonso Clementino.
Data: 30/05/2021

1. Objetivo

Observar o impacto da normalização e da escolha na quantidade de vizinhos no processo de classificação kNN (*k Nearest Neighbors*), buscando explicar os resultados encontrados.

2. Metodologia

2.1. Ambiente:

O experimento foi realizado em uma máquina local, com processador Intel Core(TM) i5-4690 CPU @ 3.50GHz, e 16GB de memória RAM.

2.2. Experimentos:

Para realização do experimento, as seguintes ações serão tomadas.

Quanto à normalização, serão gerados vetores de característica para diversas dimensões X e Y de imagens normalizadas, obtendo-se resultados para permutações de X e Y indo de 1 a 32 pixels por normalização ((1,1), (1,2), ..., (32,32)). Para cada permutação, cada uma das amostras contará com X*Y características (Quantidade de pixel). Para cada posição da imagem, verifica-se o valor de intensidade do pixel e se esse valor for > 128, a característica é igual a 1, caso contrário 0, definindo assim as características de cada amostra.

Após a geração dos vetores de características, estes passarão pelo classificador kNN, com quantidade de vizinhos (k) igual a 3 e tendo 50% de suas amostras usadas como treino e a outra metade como validação. Dado que a base é balanceada (200 amostras por classe), será observada a métrica Acurácia (*Accuracy*), definindo assim a(s) melhor(es) e pior(es) combinações XY para geração do modelo.

Além disso, após obter-se a melhor combinação XY, esta será posta a prova com múltiplas aplicações no classificador kNN, cada qual com um k (número de vizinhos) diferente, buscando identificar qual melhor k para o problema. os valores de k irão variar de 1 a 512.

2.3. Limitações:

Buscou-se expandir o domínio do experimento de variações das dimensões no processo de normalização e classificação de imagens acima do estipulado (32 pixel para cada dimensão). Porém, além do grande tempo que foi despendido na tentativa, não foi possível fazer o armazenamento de todos os dados gerados (features + resultado knn), isso implicará em alguns resultados e conclusões sobre os experimentos.

3. Resultados

A seguir serão apresentados os resultados para primeira parte do experimento “**Relação Dimensões X Y**” e para segunda parte “**Relação Quantidade de Vizinhos**”
Para melhor visualização dos resultados, é possível acessar planilhas [aqui](#).

3.1. Relação Dimensões X Y

X \ Y	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
1	0,092	0,116	0,194	0,294	0,343	0,351	0,299	0,367	0,370	0,373	0,375	0,378	0,408	0,404	0,375	0,378	0,425	0,390	0,411	0,404	0,406	0,413	0,403	0,418	0,405	0,406	0,442	0,424	0,422	0,416	0,413	0,418
2	0,143	0,217	0,350	0,437	0,477	0,539	0,509	0,528	0,536	0,554	0,574	0,575	0,555	0,582	0,581	0,607	0,590	0,591	0,588	0,604	0,580	0,594	0,607	0,597	0,583	0,622	0,604	0,594	0,586	0,616	0,607	0,595
3	0,222	0,341	0,487	0,547	0,597	0,637	0,641	0,647	0,650	0,662	0,656	0,674	0,666	0,669	0,664	0,674	0,688	0,684	0,681	0,677	0,682	0,679	0,679	0,682	0,691	0,697	0,695	0,699	0,702	0,693	0,693	0,693
4	0,286	0,458	0,594	0,683	0,733	0,773	0,770	0,784	0,790	0,799	0,790	0,803	0,803	0,817	0,814	0,815	0,812	0,821	0,818	0,829	0,828	0,830	0,828	0,821	0,816	0,815	0,829	0,812	0,813	0,821	0,819	0,819
5	0,319	0,565	0,690	0,749	0,789	0,792	0,814	0,829	0,825	0,833	0,843	0,828	0,855	0,847	0,852	0,849	0,856	0,851	0,858	0,852	0,847	0,865	0,864	0,863	0,860	0,868	0,864	0,869	0,857	0,861	0,865	0,862
6	0,341	0,579	0,715	0,793	0,826	0,853	0,843	0,852	0,851	0,864	0,874	0,873	0,878	0,869	0,876	0,876	0,877	0,879	0,879	0,883	0,887	0,884	0,883	0,885	0,888	0,884	0,875	0,875	0,885	0,879	0,883	0,883
7	0,398	0,628	0,747	0,824	0,842	0,859	0,868	0,862	0,875	0,876	0,880	0,888	0,906	0,902	0,891	0,891	0,894	0,899	0,889	0,902	0,894	0,898	0,904	0,891	0,906	0,905	0,894	0,905	0,905	0,902	0,903	0,909
8	0,467	0,634	0,767	0,829	0,835	0,855	0,872	0,864	0,885	0,879	0,891	0,889	0,888	0,884	0,895	0,894	0,892	0,898	0,888	0,898	0,892	0,893	0,897	0,890	0,896	0,902	0,896	0,897	0,896	0,894	0,890	0,896
9	0,428	0,670	0,764	0,836	0,852	0,869	0,885	0,877	0,874	0,884	0,890	0,883	0,887	0,889	0,889	0,899	0,899	0,901	0,893	0,903	0,897	0,908	0,905	0,900	0,898	0,903	0,906	0,902	0,908	0,913	0,903	0,916
10	0,424	0,674	0,795	0,836	0,871	0,877	0,877	0,884	0,883	0,894	0,892	0,904	0,906	0,901	0,900	0,901	0,903	0,901	0,904	0,905	0,899	0,903	0,912	0,903	0,914	0,901	0,902	0,914	0,904	0,907	0,900	0,902
11	0,467	0,683	0,791	0,847	0,873	0,884	0,883	0,886	0,892	0,895	0,895	0,903	0,902	0,901	0,907	0,897	0,900	0,902	0,902	0,901	0,902	0,900	0,904	0,899	0,899	0,903	0,901	0,898	0,909	0,907	0,901	0,899
12	0,450	0,679	0,799	0,863	0,872	0,886	0,885	0,901	0,887	0,893	0,905	0,897	0,909	0,913	0,903	0,905	0,908	0,909	0,912	0,917	0,916	0,916	0,913	0,907	0,911	0,919	0,911	0,909	0,912	0,905	0,912	0,912
13	0,471	0,683	0,784	0,850	0,875	0,892	0,877	0,902	0,893	0,899	0,909	0,902	0,905	0,897	0,905	0,902	0,905	0,901	0,918	0,904	0,899	0,908	0,912	0,908	0,905	0,905	0,905	0,908	0,914	0,903	0,909	0,902
14	0,472	0,704	0,811	0,861	0,870	0,887	0,891	0,905	0,897	0,893	0,909	0,904	0,903	0,908	0,904	0,908	0,901	0,911	0,910	0,914	0,917	0,910	0,908	0,909	0,911	0,910	0,909	0,912	0,912	0,916	0,913	0,913
15	0,440	0,688	0,814	0,869	0,878	0,896	0,895	0,886	0,896	0,908	0,894	0,907	0,916	0,908	0,907	0,909	0,910	0,918	0,910	0,910	0,907	0,916	0,908	0,909	0,907	0,911	0,907	0,906	0,910	0,915	0,921	0,909
16	0,470	0,702	0,817	0,861	0,887	0,895	0,896	0,900	0,902	0,909	0,912	0,903	0,910	0,916	0,907	0,914	0,911	0,915	0,912	0,919	0,919	0,923	0,911	0,923	0,924	0,925	0,919	0,918	0,924	0,925	0,921	0,919
17	0,485	0,720	0,816	0,872	0,892	0,891	0,905	0,897	0,906	0,910	0,912	0,918	0,908	0,902	0,910	0,917	0,902	0,911	0,914	0,914	0,913	0,919	0,910	0,908	0,913	0,915	0,915	0,918	0,914	0,914	0,909	0,915
18	0,483	0,683	0,807	0,860	0,881	0,901	0,899	0,896	0,897	0,902	0,906	0,906	0,900	0,906	0,908	0,910	0,906	0,908	0,903	0,909	0,911	0,916	0,908	0,913	0,913	0,912	0,908	0,913	0,909	0,912	0,914	0,910
19	0,467	0,703	0,806	0,869	0,888	0,901	0,902	0,896	0,901	0,902	0,908	0,916	0,907	0,917	0,913	0,920	0,912	0,914	0,914	0,918	0,914	0,917	0,915	0,906	0,918	0,916	0,911	0,922	0,919	0,912	0,920	0,918
20	0,470	0,685	0,816	0,867	0,883	0,905	0,904	0,899	0,906	0,904	0,910	0,910	0,910	0,908	0,903	0,918	0,914	0,915	0,921	0,922	0,914	0,918	0,915	0,913	0,919	0,918	0,917	0,920	0,920	0,918	0,919	0,914
21	0,490	0,697	0,828	0,871	0,892	0,899	0,897	0,895	0,904	0,895	0,904	0,913	0,915	0,910	0,902	0,906	0,910	0,916	0,908	0,914	0,918	0,918	0,912	0,915	0,917	0,909	0,919	0,919	0,915	0,917	0,916	0,917
22	0,486	0,701	0,817	0,861	0,881	0,898	0,909	0,894	0,907	0,905	0,912	0,913	0,911	0,902	0,906	0,908	0,909	0,909	0,909	0,914	0,916	0,918	0,914	0,910	0,912	0,916	0,906	0,913	0,915	0,914	0,911	0,910
23	0,488	0,713	0,821	0,864	0,893	0,899	0,902	0,898	0,909	0,909	0,914	0,913	0,916	0,908	0,913	0,916	0,917	0,917	0,915	0,920	0,912	0,924	0,920	0,917	0,918	0,915	0,922	0,922	0,921	0,919	0,925	0,920
24	0,490	0,704	0,818	0,865	0,884	0,901	0,894	0,900	0,908	0,908	0,908	0,912	0,907	0,916	0,907	0,913	0,915	0,915	0,917	0,920	0,916	0,919	0,912	0,913	0,912	0,919	0,915	0,917	0,913	0,913	0,915	0,917
25	0,503	0,708	0,831	0,868	0,887	0,907	0,904	0,902	0,914	0,913	0,911	0,921	0,911	0,918	0,910	0,914	0,917	0,919	0,918	0,928	0,918	0,920	0,923	0,921	0,919	0,919	0,917	0,928	0,927	0,914	0,925	0,923
26	0,492	0,710	0,837	0,860	0,882	0,910	0,909	0,904	0,909	0,909	0,911	0,913	0,915	0,911	0,916	0,916	0,913	0,920	0,915	0,922	0,917	0,917	0,922	0,920	0,918	0,919	0,915	0,923	0,922	0,918	0,921	0,918
27	0,498	0,718	0,824	0,858	0,885	0,906	0,908	0,899	0,905	0,910	0,909	0,908	0,907	0,916	0,914	0,917	0,912	0,911	0,910	0,916	0,917	0,923	0,914	0,918	0,918	0,913	0,914	0,919	0,918	0,919	0,915	0,919
28	0,509	0,719	0,822	0,874	0,889	0,909	0,898	0,903	0,902	0,911	0,913	0,917	0,909	0,913	0,915	0,914	0,913	0,916	0,918	0,919	0,911	0,919	0,919	0,920	0,921	0,914	0,919	0,923	0,918	0,918	0,917	0,917
29	0,488	0,699	0,817	0,868	0,889	0,904	0,904	0,907	0,907	0,906	0,912	0,915	0,907	0,909	0,912	0,914	0,910	0,913	0,915	0,922	0,915	0,923	0,917	0,917	0,912	0,918	0,917	0,920	0,919	0,913	0,919	0,917
30	0,504	0,719	0,840	0,875	0,894	0,897	0,904	0,910	0,916	0,908	0,919	0,908	0,906	0,909	0,911	0,914	0,912	0,915	0,912	0,920	0,914	0,914	0,916	0,919	0,919	0,916	0,912	0,921	0,915	0,916	0,916	0,919
31	0,506	0,720	0,827	0,882	0,884	0,904	0,894	0,905	0,908	0,915	0,918	0,915	0,914	0,907	0,915	0,911	0,909	0,921	0,921	0,925	0,915	0,919	0,916	0,918	0,917	0,917	0,912	0,921	0,923	0,918	0,922	0,921
32	0,519	0,714	0,837	0,870	0,903	0,908	0,902	0,904	0,910	0,905	0,913	0,913	0,909	0,908	0,913	0,909	0,911	0,914	0,919	0,915	0,915	0,918	0,913	0,915	0,913	0,911	0,909	0,915	0,915	0,915	0,920	0,915

Tabela 1: “Acurácia Para Combinação de Dimensões X (Colunas) e Y (Linhas) com K=3”

Gradiente Indica menor acurácia (células claras) para maior acurácia (células verde-escuras). A célula em borda vermelha indica Combinação X Y com pior acurácia (X=1, Y=1), enquanto as células de borda Azul-claro indicam as Combinações X Y de melhor acurácia ((X=20, Y=25) e (X=28, Y=25)).

$v \setminus x$	1	2	3	4	5	6	7	8	9	10	11	12	13	15	14	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
1	0,579	0,578	0,596	0,592	0,611	0,573	0,596	0,572	0,574	0,604	0,592	0,608	0,617	0,614	0,604	0,611	0,614	0,610	0,657	0,626	0,629	0,634	0,619	0,641	0,653	0,662	0,703	0,666	0,667	0,682	0,674	0,685
2	0,580	0,579	0,598	0,630	0,603	0,608	0,610	0,636	0,633	0,629	0,651	0,673	0,661	0,668	0,652	0,687	0,687	0,703	0,685	0,699	0,715	0,742	0,733	0,738	0,743	0,763	0,769	0,755	0,798	0,780	0,800	0,799
3	0,573	0,610	0,593	0,636	0,619	0,642	0,641	0,757	0,678	0,683	0,691	0,707	0,708	0,717	0,701	0,717	0,759	0,743	0,789	0,781	0,789	0,827	0,807	0,824	0,839	0,833	0,861	0,876	0,882	0,887	0,893	0,908
4	0,579	0,606	0,605	0,617	0,641	0,657	0,668	0,673	0,688	0,708	0,724	0,737	0,771	0,779	0,772	0,789	0,801	0,832	0,834	0,842	0,859	0,883	0,878	0,912	0,927	0,938	0,964	0,974	0,986	1,002	1,027	1,018
5	0,575	0,619	0,617	0,646	0,685	0,677	0,702	0,715	0,756	0,750	0,757	0,791	0,815	0,847	0,809	0,868	0,879	0,918	0,928	0,927	0,957	0,959	0,982	1,014	1,019	1,068	1,069	1,095	1,130	1,176	1,141	1,175
6	0,571	0,629	0,663	0,693	0,681	0,693	0,732	0,736	0,761	0,786	0,809	0,819	0,885	0,879	0,878	0,932	0,947	0,983	0,994	0,987	1,030	1,067	1,082	1,097	1,141	1,155	1,173	1,192	1,238	1,262	1,279	1,268
7	0,580	0,622	0,670	0,704	0,701	0,711	0,739	0,759	0,813	0,824	0,856	0,885	0,889	0,940	0,906	0,944	0,982	1,051	1,047	1,076	1,100	1,140	1,139	1,175	1,218	1,251	1,270	1,302	1,329	1,451	1,382	1,395
8	0,593	0,623	0,649	0,707	0,728	0,754	0,790	0,796	0,833	0,859	0,888	0,927	0,965	1,005	0,969	1,028	1,060	1,072	1,126	1,151	1,163	1,212	1,239	1,257	1,290	1,333	1,355	1,398	1,414	1,450	1,504	1,509
9	0,603	0,636	0,679	0,680	0,734	0,788	0,805	0,809	0,844	0,900	0,947	0,947	1,007	1,089	1,022	1,109	1,146	1,179	1,247	1,226	1,299	1,353	1,404	1,407	1,435	1,515	1,574	1,580	1,601	1,685	1,675	1,789
10	0,601	0,654	0,728	0,713	0,759	0,811	0,837	0,862	0,884	0,945	0,992	1,007	1,058	1,151	1,084	1,152	1,219	1,266	1,247	1,359	1,383	1,426	1,475	1,485	1,545	1,618	1,673	1,673	1,700	1,791	1,777	1,883
11	0,613	0,655	0,693	0,739	0,785	0,799	0,845	0,896	0,926	0,986	1,038	1,043	1,112	1,185	1,155	1,226	1,249	1,339	1,391	1,416	1,436	1,515	1,534	1,565	1,638	1,690	1,761	1,771	1,818	1,948	1,905	2,018
12	0,609	0,666	0,742	0,733	0,787	0,822	0,883	0,908	0,974	0,986	1,037	1,082	1,154	1,243	1,187	1,262	1,353	1,378	1,444	1,476	1,535	1,577	1,614	1,662	1,704	1,803	1,871	1,856	1,918	2,015	2,019	2,130
13	0,572	0,677	0,725	0,752	0,818	0,862	0,892	0,968	1,024	1,078	1,118	1,163	1,202	1,329	1,264	1,391	1,459	1,474	1,528	1,648	1,654	1,712	1,827	1,824	1,983	1,947	2,104	2,172	2,119	2,285	2,262	2,316
14	0,612	0,675	0,713	0,780	0,812	0,885	0,930	1,024	1,016	1,093	1,147	1,191	1,260	1,388	1,307	1,452	1,540	1,568	1,608	1,703	1,720	1,811	1,905	1,934	2,078	2,034	2,203	2,277	2,239	2,431	2,382	2,453
15	0,604	0,677	0,729	0,801	0,859	0,908	0,943	1,022	1,064	1,150	1,211	1,264	1,298	1,411	1,336	1,500	1,586	1,626	1,691	1,766	1,793	1,890	2,001	2,031	2,145	2,134	2,286	2,386	2,340	2,535	2,517	2,559
16	0,626	0,686	0,757	0,794	0,864	0,922	0,986	1,069	1,118	1,155	1,226	1,307	1,336	1,468	1,401	1,564	1,664	1,675	1,770	1,849	1,872	1,951	2,114	2,093	2,275	2,258	2,394	2,484	2,446	2,595	2,596	2,695
17	0,630	0,687	0,761	0,842	0,903	0,948	1,035	1,096	1,149	1,234	1,282	1,371	1,455	1,622	1,507	1,710	1,797	1,808	1,882	1,947	1,995	2,088	2,323	2,278	2,463	2,504	2,539	2,595	2,765	2,832	2,972	3,004
18	0,664	0,682	0,761	0,860	0,907	1,038	1,040	1,107	1,198	1,243	1,336	1,401	1,515	1,679	1,555	1,738	1,909	1,843	1,932	2,046	2,161	2,203	2,408	2,418	2,556	2,642	2,621	2,721	2,864	2,969	3,027	3,105
19	0,695	0,705	0,766	0,844	0,931	0,979	1,065	1,132	1,192	1,287	1,399	1,459	1,536	1,719	1,613	1,811	1,906	1,940	2,024	2,109	2,223	2,290	2,456	2,444	2,677	2,764	2,766	2,828	2,988	3,064	3,206	3,235
20	0,611	0,714	0,803	0,877	0,940	1,014	1,114	1,164	1,250	1,318	1,426	1,499	1,587	1,801	1,664	1,870	1,983	2,012	2,105	2,183	2,307	2,372	2,584	2,536	2,710	2,811	2,827	2,924	3,093	3,174	3,263	3,340
21	0,693	0,715	0,816	0,880	0,945	1,084	1,161	1,197	1,311	1,379	1,484	1,626	1,653	1,844	1,769	1,946	2,037	2,237	2,244	2,318	2,425	2,642	2,744	2,746	2,833	2,938	3,106	3,154	3,314	3,363	3,437	3,573
22	0,690	0,725	0,816	0,900	1,000	1,096	1,179	1,279	1,312	1,422	1,528	1,687	1,715	1,911	1,818	2,007	2,078	2,288	2,296	2,397	2,544	2,627	2,905	2,864	2,929	3,043	3,190	3,225	3,423	3,512	3,551	3,705
23	0,790	0,744	0,863	0,893	0,997	1,057	1,172	1,272	1,377	1,545	1,544	1,707	1,755	1,964	1,870	2,116	2,180	2,364	2,369	2,506	2,642	2,831	2,919	2,945	3,015	3,136	3,314	3,361	3,499	3,620	3,654	3,806
24	0,670	0,764	0,841	0,916	1,018	1,136	1,218	1,303	1,415	1,524	1,617	1,723	1,836	2,001	1,974	2,080	2,222	2,435	2,463	2,561	2,697	2,871	2,983	3,076	3,131	3,222	3,414	3,499	3,600	3,737	3,811	3,937
25	0,686	0,744	0,853	1,080	1,042	1,132	1,269	1,329	1,471	1,614	1,655	1,792	1,903	2,210	1,993	2,392	2,467	2,481	2,615	2,755	2,887	2,965	3,106	3,225	3,360	3,424	3,629	3,702	3,888	3,898	4,007	4,106
26	0,718	0,756	0,856	1,076	1,060	1,158	1,262	1,375	1,469	1,632	1,709	1,834	1,942	2,268	2,071	2,447	2,537	2,541	2,694	2,802	2,962	3,054	3,175	3,328	3,489	3,525	3,706	3,790	3,947	4,013	4,199	4,247
27	0,719	0,788	0,880	1,004	1,044	1,183	1,299	1,382	1,510	1,684	1,778	1,868	1,997	2,329	2,125	2,490	2,590	2,623	2,751	2,900	3,041	3,150	3,273	3,373	3,583	3,641	3,806	3,913	4,031	4,143	4,298	4,424
28	0,719	0,765	0,880	0,929	1,085	1,205	1,326	1,434	1,566	1,728	1,787	1,940	2,027	2,384	2,160	2,519	2,651	2,685	2,843	2,956	3,078	3,230	3,322	3,468	3,640	3,715	3,869	4,002	4,134	4,212	4,390	4,516
29	0,663	0,798	0,906	1,047	1,123	1,241	1,372	1,507	1,644	1,739	1,952	2,066	2,143	2,528	2,398	2,636	2,720	2,791	3,030	3,197	3,213	3,384	3,475	3,632	3,825	3,920	4,110	4,243	4,334	4,433	4,530	4,751
30	0,686	0,800	0,923	1,051	1,151	1,278	1,380	1,542	1,669	1,792	2,008	2,115	2,176	2,562	2,341	2,690	2,802	2,910	3,091	3,228	3,331	3,479	3,608	3,750	3,929	4,031	4,143	4,323	4,393	4,559	4,692	4,839
31	0,677	0,795	0,900	1,067	1,173	1,278	1,422	1,515	1,724	1,816	2,025	2,149	2,224	2,645	2,485	2,764	2,796	2,939	3,169	3,308	3,425	3,571	3,703	3,844	4,008	4,142	4,288	4,428	4,586	4,695	4,809	4,969
32	0,687	0,820	0,922	1,089	1,180	1,313	1,451	1,611	1,750	1,821	2,067	2,209	2,271	2,682	2,530	2,863	2,879	3,021	3,267	3,376	3,505	3,658	3,774	3,942	4,097	4,234	4,363	4,521	4,633	4,757	5,058	5,037

Tabela 2: “Tempo de Processo Completo de Classificador KNN para Combinação de Dimensões X (Colunas) e Y (Linhas) com K=3”

Tabela expressa graduação no aumento de tempo de acordo com a aplicação do gradiente (branco menor tempo, vermelho-escuro maior tempo) em relação ao tamanho das imagens / quantidade de atributos da amostra, indicados por X (Coluna) e Y (Linha).

$V \setminus X$	1	2	3	4	5	6	7	8	9	10	11	12	13	15	14	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
1	0,159	0,201	0,326	0,497	0,561	0,613	0,502	0,642	0,645	0,618	0,633	0,622	0,661	0,658	0,621	0,619	0,692	0,639	0,626	0,645	0,645	0,651	0,651	0,652	0,620	0,613	0,629	0,637	0,633	0,610	0,613	0,610
2	0,247	0,375	0,585	0,694	0,791	0,887	0,834	0,830	0,847	0,881	0,882	0,854	0,840	0,871	0,891	0,884	0,859	0,841	0,858	0,864	0,811	0,801	0,828	0,809	0,785	0,815	0,785	0,787	0,734	0,790	0,759	0,745
3	0,387	0,559	0,821	0,860	0,964	0,992	1,000	0,855	0,959	0,969	0,949	0,953	0,941	0,933	0,947	0,940	0,906	0,921	0,863	0,867	0,864	0,821	0,841	0,828	0,824	0,837	0,807	0,798	0,796	0,781	0,776	0,763
4	0,494	0,756	0,982	1,107	1,144	1,177	1,153	1,165	1,148	1,129	1,091	1,090	1,042	1,049	1,054	1,033	1,014	0,987	0,981	0,985	0,964	0,940	0,943	0,900	0,880	0,869	0,860	0,834	0,825	0,819	0,797	0,805
5	0,555	0,913	1,118	1,159	1,152	1,170	1,160	1,159	1,091	1,111	1,114	1,047	1,049	1,000	1,053	0,978	0,974	0,927	0,925	0,919	0,885	0,902	0,880	0,851	0,844	0,813	0,808	0,794	0,758	0,732	0,758	0,734
6	0,597	0,921	1,078	1,144	1,213	1,231	1,152	1,158	1,118	1,099	1,080	1,066	0,992	0,989	0,998	0,940	0,926	0,894	0,884	0,895	0,861	0,828	0,816	0,807	0,778	0,765	0,746	0,734	0,715	0,697	0,690	0,696
7	0,686	1,010	1,115	1,170	1,201	1,208	1,175	1,136	1,076	1,063	1,028	1,003	1,019	0,960	0,983	0,944	0,910	0,855	0,849	0,838	0,813	0,788	0,794	0,758	0,744	0,723	0,704	0,695	0,681	0,622	0,653	0,652
8	0,788	1,018	1,182	1,173	1,147	1,134	1,104	1,085	1,062	1,023	1,003	0,959	0,920	0,880	0,924	0,870	0,842	0,838	0,789	0,780	0,767	0,737	0,724	0,708	0,695	0,677	0,661	0,642	0,634	0,617	0,592	0,594
9	0,710	1,053	1,125	1,229	1,161	1,103	1,099	1,084	1,036	0,982	0,940	0,932	0,881	0,816	0,870	0,811	0,784	0,764	0,716	0,737	0,691	0,671	0,645	0,640	0,626	0,596	0,576	0,571	0,567	0,542	0,539	0,512
10	0,705	1,031	1,092	1,173	1,148	1,081	1,048	1,026	0,999	0,946	0,899	0,898	0,856	0,783	0,830	0,782	0,741	0,712	0,725	0,666	0,650	0,633	0,618	0,608	0,592	0,557	0,539	0,546	0,532	0,506	0,506	0,479
11	0,762	1,043	1,141	1,146	1,112	1,106	1,045	0,989	0,963	0,908	0,862	0,866	0,811	0,760	0,785	0,732	0,721	0,674	0,648	0,636	0,628	0,594	0,589	0,574	0,549	0,534	0,512	0,507	0,500	0,466	0,473	0,445
12	0,739	1,020	1,077	1,177	1,108	1,078	1,002	0,992	0,911	0,906	0,873	0,829	0,788	0,735	0,761	0,717	0,671	0,660	0,632	0,621	0,597	0,581	0,566	0,546	0,535	0,510	0,487	0,490	0,475	0,449	0,452	0,428
13	0,823	1,009	1,081	1,130	1,070	1,035	0,983	0,932	0,872	0,834	0,813	0,776	0,753	0,675	0,716	0,648	0,620	0,611	0,601	0,549	0,544	0,530	0,499	0,498	0,456	0,465	0,430	0,418	0,431	0,395	0,402	0,389
14	0,771	1,043	1,137	1,104	1,071	1,002	0,958	0,884	0,883	0,817	0,793	0,759	0,717	0,654	0,692	0,625	0,585	0,581	0,566	0,537	0,533	0,502	0,477	0,470	0,438	0,447	0,413	0,401	0,407	0,377	0,383	0,372
15	0,728	1,016	1,117	1,085	1,022	0,987	0,949	0,867	0,842	0,790	0,738	0,718	0,706	0,644	0,679	0,606	0,574	0,565	0,538	0,515	0,506	0,485	0,454	0,448	0,423	0,427	0,397	0,380	0,389	0,361	0,366	0,355
16	0,751	1,023	1,079	1,084	1,027	0,971	0,909	0,842	0,807	0,787	0,744	0,691	0,681	0,624	0,647	0,584	0,547	0,546	0,515	0,497	0,491	0,473	0,431	0,441	0,406	0,410	0,384	0,370	0,378	0,356	0,355	0,341
17	0,770	1,048	1,072	1,036	0,988	0,940	0,874	0,818	0,789	0,737	0,711	0,670	0,624	0,556	0,604	0,536	0,502	0,504	0,486	0,469	0,458	0,440	0,392	0,399	0,371	0,365	0,360	0,354	0,331	0,323	0,306	0,305
18	0,727	1,001	1,060	1,000	0,971	0,868	0,864	0,809	0,749	0,726	0,678	0,647	0,594	0,540	0,584	0,524	0,475	0,493	0,467	0,444	0,422	0,416	0,377	0,378	0,357	0,345	0,346	0,336	0,317	0,307	0,302	0,293
19	0,672	0,997	1,052	1,030	0,954	0,920	0,847	0,792	0,756	0,701	0,649	0,628	0,590	0,533	0,566	0,508	0,478	0,471	0,452	0,435	0,411	0,400	0,373	0,371	0,343	0,331	0,329	0,326	0,308	0,298	0,287	0,284
20	0,769	0,959	1,016	0,989	0,939	0,893	0,811	0,772	0,725	0,686	0,638	0,607	0,573	0,504	0,543	0,491	0,461	0,455	0,438	0,422	0,396	0,387	0,354	0,360	0,339	0,327	0,324	0,315	0,297	0,289	0,282	0,274
21	0,707	0,975	1,015	0,990	0,944	0,829	0,773	0,748	0,690	0,649	0,609	0,562	0,554	0,493	0,510	0,466	0,447	0,409	0,405	0,394	0,379	0,347	0,332	0,333	0,324	0,309	0,296	0,291	0,276	0,273	0,267	0,257
22	0,704	0,967	1,001	0,957	0,881	0,819	0,771	0,699	0,691	0,636	0,597	0,541	0,531	0,472	0,498	0,452	0,437	0,397	0,396	0,381	0,360	0,349	0,315	0,318	0,311	0,301	0,284	0,283	0,267	0,260	0,257	0,246
23	0,618	0,958	0,951	0,968	0,896	0,851	0,770	0,706	0,660	0,588	0,592	0,535	0,522	0,462	0,488	0,433	0,421	0,388	0,386	0,367	0,345	0,326	0,315	0,311	0,304	0,292	0,278	0,274	0,263	0,254	0,253	0,242
24	0,731	0,921	0,973	0,944	0,868	0,793	0,734	0,691	0,642	0,596	0,562	0,529	0,494	0,458	0,459	0,439	0,412	0,376	0,372	0,359	0,340	0,320	0,306	0,297	0,291	0,285	0,268	0,262	0,254	0,244	0,240	0,233
25	0,733	0,952	0,974	0,804	0,851	0,801	0,712	0,679	0,621	0,566	0,550	0,514	0,479	0,415	0,457	0,382	0,372	0,370	0,351	0,337	0,318	0,310	0,297	0,286	0,274	0,268	0,253	0,251	0,238	0,234	0,231	0,225
26	0,685	0,939	0,978	0,799	0,832	0,786	0,720	0,657	0,619	0,557	0,533	0,498	0,471	0,402	0,442	0,374	0,360	0,362	0,340	0,329	0,310	0,300	0,290	0,276	0,263	0,261	0,247	0,244	0,234	0,229	0,219	0,216
27	0,693	0,911	0,936	0,855	0,848	0,766	0,699	0,651	0,599	0,540	0,511	0,486	0,454	0,393	0,430	0,368	0,352	0,347	0,331	0,316	0,302	0,293	0,279	0,272	0,256	0,251	0,240	0,235	0,228	0,222	0,213	0,208
28	0,708	0,940	0,934	0,941	0,819	0,754	0,677	0,630	0,576	0,527	0,511	0,473	0,448	0,383	0,424	0,363	0,344	0,341	0,323	0,311	0,296	0,285	0,277	0,265	0,253	0,246	0,238	0,231	0,222	0,218	0,209	0,203
29	0,736	0,876	0,902	0,829	0,792	0,728	0,659	0,602	0,552	0,521	0,467	0,443	0,423	0,360	0,380	0,347	0,335	0,327	0,302	0,288	0,285	0,273	0,264	0,252	0,238	0,234	0,223	0,217	0,212	0,206	0,203	0,193
30	0,735	0,899	0,910	0,833	0,777	0,702	0,655	0,590	0,549	0,507	0,458	0,429	0,416	0,355	0,389	0,340	0,325	0,314	0,295	0,285	0,274	0,263	0,254	0,245	0,234	0,227	0,220	0,213	0,208	0,201	0,195	0,190
31	0,747	0,906	0,919	0,827	0,754	0,707	0,629	0,597	0,527	0,504	0,453	0,426	0,411	0,343	0,368	0,330	0,325	0,313	0,291	0,280	0,267	0,257	0,247	0,239	0,229	0,221	0,213	0,208	0,201	0,196	0,192	0,185
32	0,755	0,871	0,908	0,799	0,765	0,692	0,622	0,561	0,520	0,497	0,442	0,413	0,400	0,339	0,361	0,317	0,316	0,303	0,281	0,271	0,261	0,251	0,242	0,232	0,223	0,215	0,208	0,202	0,197	0,192	0,182	0,182

Tabela 3: “Relação Acurácia / Tempo para Cada Processo Completo de Classificador KNN para Cada Combinação de Dimensões X (Colunas) e Y (Linhas) com K=3”

Tabela mostra proporção Acurácia/Tempo (Indicados em **Tabela 1** e **Tabela 2**). Pontos brancos possuem proporção baixa (pior relação custo desempenho) enquanto pontos azul-escuros indicam proporção alta (melhor relação custo desempenho).

3.2. Relação Quantidade de Vizinhos

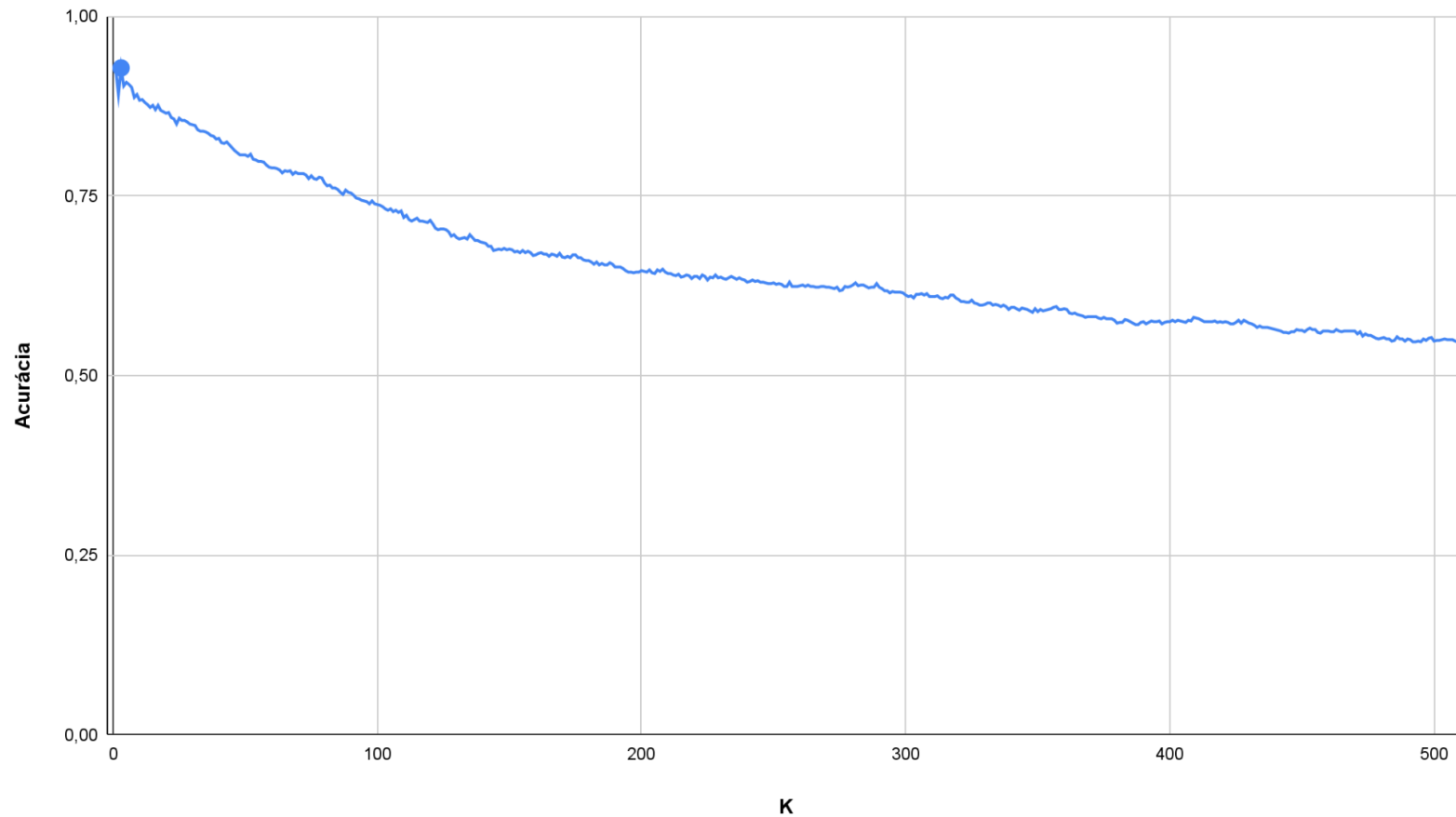


Gráfico 1: “Acurácia Variando-se K, com Imagens Normalizadas X=20, Y=25”

Gráfico indica Acurácia obtida para os experimentos contendo normalização de imagens fixa X=20 e Y=25 (500 atributos por amostra). variando-se quantidade de vizinhos para classificação (de 1 à 512). Em círculo azul, destaca-se a melhor acurácia (0,928) para K = 3.

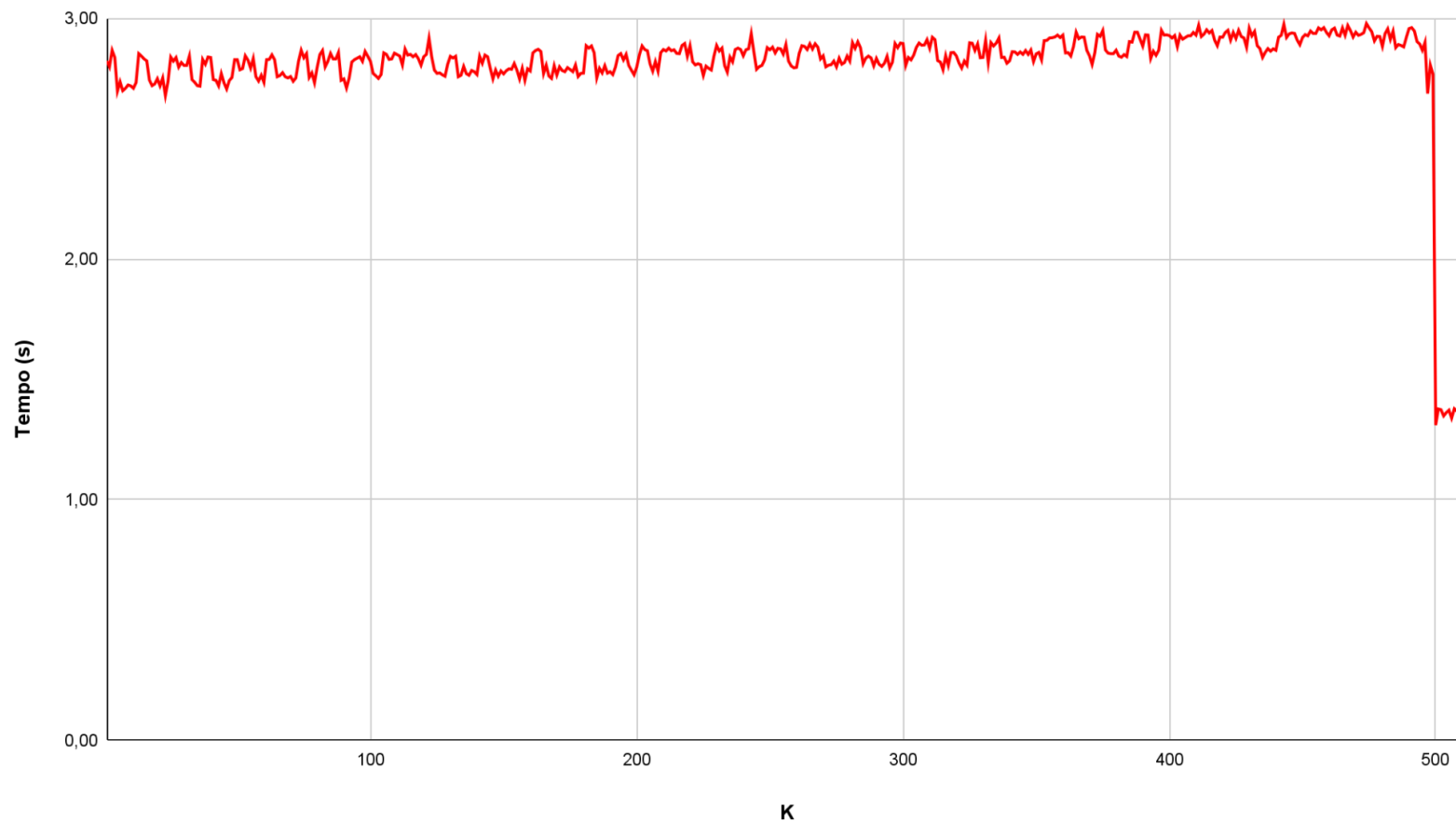


Gráfico 1: “Acurácia Variando-se K, com Imagens Normalizadas X=20, Y=25”

Tempo gasto para os experimentos contendo normalização de imagens fixa X=20 e Y=25 (500 atributos por amostra), variando-se a quantidade de vizinhos para classificação (de 1 à 512). Destaca-se que a partir de K = 500 uma diminuição considerável no tempo de treino e avaliação.

4. Conclusão

Observando a **Tabela 1** é possível notar que, para o domínio estudado no experimento (X e Y variando de 1 à 32), o acréscimo de pixel e, consequentemente, expansão no número de atributos e dimensão do problema, foi em boa parte benéfico no quesito acurácia do modelo. Dadas as limitações, tal resultado vai em contrapartida do estudado em classe quanto a expansão no número de atributos prejudica a acurácia do modelo. Tal comportamento talvez pudesse ser observado caso ampliado o domínio estudado (problema abordado em **2.3. Limitações**).

Para a **Tabela 2** é possível observar o impacto do recurso tempo no período de treino e avaliação de um modelo conforme o ocorre o acréscimo de atributos no modelo estudado.

Quanto a **Tabela 3**, operasse, a partir dos resultados das **Tabela 1** e **Tabela 2**, a relação **Acurácia/Tempo**, dessa forma é possível identificar os pontos nos quais ocorre a melhor utilização do tempo em relação ao ganho em acurácia (tons de azul mais escuro), ou seja, em cenários de recurso tempo limitado, quais seriam as melhores dimensões de custo-benefício para a execução do estudo.

Para os **Gráficos 1 e 2**, nota-se que as melhores escolhas de vizinhos são as menores, com K=3 retornando acurácia de 0,928, seguido por K=1, com acurácia 0,925.

Além disso, nota-se que quanto maior a quantidade de vizinhos estudada pior se torna a classificação. A hipótese explica tal comportamento é a de que o problema estudado possui múltiplas fronteiras, dessa forma os elementos de uma classe estariam dispersos no espaço em pequenos grupos, dificultando a classificação de um novo elemento de determinada classe conforme a quantidade de vizinhos é aumentada, tal problema pode ser provocado pelo método de obtenção de vetor de características das amostras ou da própria natureza dos dados de entrada.

Outro ponto a se destacar é o tempo gasto nas operações conforme o incremento de K. Como esperado, quanto maior a quantidade de vizinhos a serem comparados, maior o tempo despendido no processo, porém, como pode-se notar no **Gráfico 2**, entre K=499 e K=500 existe uma grande queda no tempo despendido. Mais de um experimento variando K entre 1 e 512 foi realizado e o mesmo comportamento foi apresentado. Não foi encontrado motivo para tal decréscimo de tempo.