Appendix A

Definition of Mathematical Notation

Notation	**Definition**
N	Number of data-points in dataset
D	Dimensionality of data-point feature representation
K	Number of hashcode bits
L	Number of hashtables
Q	Number of query data-points
B	Number bits per projected dimension
T	Number of thresholds per projected dimension
M	Iterations
C	Randomly sampled data-points *or* cluster centroids
$\mathbf{X} \in \mathbb{R}^{N \times D}$	Dataset of N data-points, dimensionality D
$\mathbf{x}_r \in \mathbb{R}^D : \mathbf{x}_r = \mathbf{X}_{r\bullet}$	r^{th} row of matrix $\mathbf{X}$
$\mathbf{x}^c \in \mathbb{R}^N : \mathbf{x}^c = \mathbf{X}_{\bullet c}$	c^{th} column of matrix $\mathbf{X}$
X_{ij}	Element of matrix $\mathbf{X}$ in row i column j
$\mathbf{q} \in \mathbb{R}^D$	Query data-point
$\mathbf{p} \in \mathbb{R}^D$	Arbitrary database data-point
$\mathbf{Y} \in \mathbb{R}^{N \times K}$	Projection matrix of N data-points, dimensionality K
$\mathbf{y}_r \in \mathbb{R}^K : \mathbf{y}_r = \mathbf{Y}_{r\bullet}$	Projected values for r^{th} data-point
$\mathbf{y}^c \in \mathbb{R}^N : \mathbf{y}^c = \mathbf{Y}_{\bullet c}$	c^{th} projected dimension
$d : \mathbb{R}^D \times \mathbb{R}^D \rightarrow \mathbb{R}$	Distance function e.g. Euclidean distance
$q_k : \mathbb{R} \rightarrow \{0,1\}^B$	Quantisation function
$\mathbf{D} \in \mathbb{R}^{N \times N}$	Matrix of data-point distances
$\mathbf{B} \in \{-1,1\}^{N \times K}$	Hashcodes of N data-points each of length K bits
$\mathbf{b}_r \in \{-1,1\}^K : \mathbf{b}_r = \mathbf{B}_{r\bullet}$	Hashcode of r^{th} data-point $\mathbf{x}_r$

$h_k : \mathbb{R}^D \rightarrow \{0,1\}$	Hash function
$g_l : \mathbb{R}^D \rightarrow \{0,1\}^K$	Hash function concatenation $[h_1(.), h_2(.), \ldots, h_K(.)]$
$\mathbf{S} \in \mathbb{R}^{N \times N}$	$S_{ij} = 1$ if $\mathbf{x}_i$ and $\mathbf{x}_j$ are nearest neighbours, 0 otherwise
$\mathbf{h}_k \in \mathbb{R}^D$	Hyperplane
$\mathbf{w}_k \in \mathbb{R}^D$	Hyperplane normal vector
$\mathbf{W} \in \mathbb{R}^{D \times K}$	Matrix of K hyperplane normal vectors
$t_k \in \mathbb{R}$	Scalar threshold
$\mathbf{T} \in \mathbb{R}^{K \times T}$	Matrix of thresholds for each projected dimension
$\mathbf{t}_r \in \mathbb{R}^T : \mathbf{t}_r = \mathbf{T}_{r\bullet}$	Set of thresholds for r^{th} projected dimension
$\kappa : \mathbb{R}^D \times \mathbb{R}^D \rightarrow \mathbb{R}$	Kernel function
$\gamma \in \mathbb{R}$	Kernel bandwidth parameter
$\|\mathbf{X}\|_F^2 = \sum_{ij}^N \|X_{ij}\|^2$	Frobenius L_2 norm of matrix
$\|\mathbf{X}\|_F^1 = \sum_{ij}^N \|X_{ij}\|$	Frobenius L_1 norm of matrix
$\mathbf{X} = diag(\mathbf{x})$	Places elements of vector $\mathbf{x}$ on diagonal of matrix $\mathbf{X}$
$sgn(a) \in \{-1,1\}$	Sign function returning 1 for $a > 0$, and -1 otherwise
$[a]_+$	Equal to a if $a \geq 0$, and 0 otherwise

Table A.1: Definition of the mathematical notation used throughout the thesis