

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



دانشگاه صنعتی امیرکبیر

(پلی تکنیک تهران)

دانشکده مهندسی صنایع

پایان نامه کارشناسی ارشد

عنوان پایان نامه:

بکارگیری تکنیک‌های داده کاوی در تعیین میزان انسولین برای
بیماران دیابتی

نگارش:

محمد سعید زائری

استاد راهنما:

آقای دکتر جمال شهرابی

بهمن ماه ۱۳۸۶

بسمه تعالی



دانشگاه صنعتی امیرکبیر
(پلی تکنیک تهران)

معاونت پژوهشی

فرم اطلاعات پایان نامه
کارشناسی ارشد و دکترا

تاریخ:

پیوست:

نام و نام خانوادگی:	محمد سعید زائری	دانشجوی آزاد	<input type="checkbox"/> بورسیه	<input type="checkbox"/> معادل	
شماره دانشجویی:	۸۴۱۲۵۰۲۰	دانشکده:	مهندسی صنایع	رشته تحصیلی:	صنایع - صنایع
نام و نام خانوادگی استاد راهنما:	جمال شهرابی				
عنوان پایان نامه به فارسی:	بکارگیری تکنیک‌های داده کاوی در تعیین میزان انسولین برای بیماران دیابتی				
عنوان پایان نامه به انگلیسی:	Applicability of Data Mining techniques to Determine the Insulin Dose amount for Diabetic Patients				
نوع پروژه:	<input checked="" type="checkbox"/> کارشناسی ارشد <input type="checkbox"/> دکترا	<input checked="" type="checkbox"/> کاربردی	<input type="checkbox"/> بنیادی	<input type="checkbox"/> توسعه ای	<input type="checkbox"/> نظری
تاریخ شروع:	مهر ۸۵	تاریخ خاتمه:	بهمن ۸۶	تعداد واحد:	۶
سازمان تأمین کننده اعتبار:	دانشگاه صنعتی امیرکبیر				
واژه های کلیدی به فارسی:	داده کاوی، تجزیه و تحلیل چند متغیره، دیابت، انسولین				
واژه های کلیدی به انگلیسی:	Data Mining, Multivariate Data Analysis, Diabetes, Insulin				
نظرها و پیشنهادهای به منظور بهبود فعالیت های پژوهشی دانشگاه:	استاد راهنما:				
دانشجو:	امضاء استاد راهنما:				
تاریخ:	تاریخ:				
نسخه ۱: معاونت پژوهشی	نسخه ۲: کتابخانه و به انضمام دو جلد پایان نامه به منظور تسویه حساب با کتابخانه و مرکز اسناد و مدارک علمی				

تقدیم به پدر و مادرم

که پیام آوران عشق، ایمان، خلوص و فداکاریند

و

تقدیم به همسرم

که پیام آور صبر و مهربانیست

تقدیر و سپاس

بدینوسیله مراتب احترام و تشکر خود را از تمامی عزیزانی که مرا در تهیه این پایان نامه یاری کرده‌اند، به ویژه **جناب آقای دکتر شهبابی** که هدایت و راهنمایی پایان نامه را بر عهده داشته‌اند اعلام می‌دارم.

همچنین از اساتید بزرگوار آقایان **دکتر قدسی پور** و **دکتر اکبرپور** که داوری این رساله را تقبل فرمودند سپاسگزارم و از خداوند متعال توفیق روزافزون آنها را خواستارم.

محمد سعید زائری

بهمن ۱۳۸۶

چکیده:

پیشرفت‌های اخیر علوم مختلف در زمینه کسب، ذخیره و نگهداری داده‌ها منجر به افزایش بسیار زیادی در حجم و ابعاد پایگاه داده‌ها شده است. همیشه در این نوع پایگاه داده‌ها، اطلاعاتی جهت استخراج و استفاده وجود دارد. صنایع پزشکی، جزء پنج صنعتی می‌باشد که با حجم زیاد داده و اطلاعات درگیر هستند. اطلاعات، داده‌ها و دانش ذخیره شده در این صنایع هر روز بسیار افزایش می‌یابد. داده کاوی یک رشته علمی جدید در زمینه بازاریابی و استخراج اطلاعات است که اقدام به استخراج دانش و ویژگی‌های مفید و جالب از مجموعه پایگاه داده‌ها می‌کند. به عبارت دیگر داده کاوی به عنوان یک قدم از مراحل استخراج دانش از پایگاه داده‌ها مورد استفاده قرار می‌گیرد و به رشته‌های آمار، **Machine Learning**، شناسایی الگوها و دسته‌بندی مشاهدات متصل شده است. امروزه از دیابت به عنوان سومین دلیل مرگ و میر دنیا یاد می‌کنند. در هر ۲۰ ثانیه یک نفر به علت ابتلاء به دیابت جان خود را از دست می‌دهد. به همین دلیل کشف و استخراج الگوها و ساختار نهان در بین معیارها و عوامل موثر بر دیابت در بین بیماران دیابتی امری ضروری و حیاتی است.

به این منظور این پایان‌نامه یک متدولوژی جدید که شامل دو قدم کلی است را معرفی می‌-

کند. در اولین قدم با بررسی سه الگوریتم **Random Projection, Non - Negative Matrix Factorization** و **PCA** اقدام به ارزیابی آنها با پیاده‌سازی آنها بر روی سه دسته داده مجازی با ساختار خطی، غیرخطی و متنی می‌کنند. نتایج بدست آمده نشان می‌دهد که **PCA** در مقایسه با دو الگوریتم دیگر توانایی بهتری در ارائه مجدد داده‌ها و کاهش ابعاد دارد. در دوم این متدولوژی در ابتدا با استفاده از تکنیک آنالیز فاکتور اقدام به استخراج الگو و ساختار در بین متغیرهای تاثیرگذار بر مسئله می‌کنند سپس با پیاده‌سازی تکنیک **PCA** به عنوان یک تکنیک کاهش ابعاد ویژگی‌های اصلی مسئله انتخاب و کاهش ابعاد انجام می‌شود. و در نهایت این متدولوژی با اجرای تکنیک آنالیز خوشه-بندی اقدام به کلاسه‌بندی بیماران می‌کنند. نتایج بدست آمده نشان می‌دهد که با کمک این متدولوژی پس از کشف الگوها و روابط بین متغیرها تنها با لحاظ کردن ۹ متغیر می‌توان مسئله را مورد بررسی قرار داده و بیماران را در جهت تعیین میزان انسولین مورد نیاز دسته‌بندی نمود.

واژگان کلیدی:

داده کاوی - تجزیه و تحلیل چند متغیره - دیابت - انسولین

فهرست

صفحه	عنوان
۱۳	فهرست اشکال
۱۵	فهرست جداول
	فصل اول : کلیات طرح تحقیق
۱۸	۱-۱- مقدمه
۱۸	۲-۱- ضرورت تحقیق
۱۹	۳-۱- بیان مسئله
۲۱	۱-۳-۱- محدودیت‌های مسئله
۲۴	۴-۱- داده کاوی
۲۶	۱-۴-۱- نگاهی گذرا به تکنیک های چند متغیره مورد استفاده
۲۶	۱-۴-۱-۱- تحلیل اجزاء اصلی
۲۶	۱-۴-۱-۲- تحلیل فاکتور
۲۶	۱-۴-۱-۳- تحلیل خوشه‌بندی
۲۶	۵-۱- متدولوژی پیشنهادی
۲۷	۶-۱- اهداف تحقیق
۲۸	۷-۱- دستاوردها
	فصل دوم: دیابت
۳۰	۱-۲- مقدمه

۳۰ ۲-۲- واکنش بدن در مقابل گلوکز
۳۲ ۱-۲-۲- سطح قند خون بالا
۳۳ ۳-۲- انواع دیابت
۳۳ ۱-۳-۲- دیابت نوع ۱
۳۳ ۱-۱-۳-۲- لوزالمعده
۳۴ ۲-۳-۲- دیابت نوع ۲
۳۵ ۳-۳-۲- دیابت دوران بلوغ
۳۵ ۴-۳-۲- دیابت حاملگی
۳۶ ۴-۲- تشخیص دیابت
۳۶ ۱-۴-۲- آزمایش HbA1c
۳۶ ۵-۲- عوارض دیابت
۳۷ ۶-۲- تغذیه و دیابت
۳۷ ۱-۶-۲- گروه‌های غذایی
۳۹ ۷-۲- انسولین
۴۱ ۸-۲- خطرات عمده دیابت
۴۱ ۱-۸-۲- هیپوگلیسمی
۴۲ ۲-۸-۲- هیپرگلیسمی

فصل سوم: تکنیک های مورد استفاده

۴۴ ۱-۳- آنالیز فاکتور
۴۴ ۱-۱-۳- مقدمه
۴۵ ۲-۱-۳- تاریخچه آنالیز فاکتور

۴۷ ۳-۱-۳- اصول تحلیلی فاکتور
۴۷ ۱-۳-۱-۳- مدل سازی تحلیل فاکتور
۴۸ ۲-۳-۱-۳- ماتریس فاکتور
۴۸ ۳-۳-۱-۳- نمره‌های فاکتورها
۴۹ ۴-۳-۱-۳- فرضیات در آنالیز فاکتور
۴۹ ۵-۳-۱-۳- تست‌های آماری
۵۰ ۶-۳-۱-۳- معیار برای تعداد استخراج فاکتور
۵۱ ۲-۳- آنالیز خوشه‌بندی
۵۱ ۱-۲-۳- مقدمه
۵۱ ۲-۲-۳- مفهوم آنالیز خوشه‌بندی
۵۲ ۳-۲-۳- اهداف آنالیز خوشه‌بندی
۵۳ ۴-۲-۳- انتخاب متغیرهای خوشه‌بندی
۵۴ ۵-۲-۳- طرح تحقیق در آنالیز خوشه‌بندی
۵۴ ۶-۲-۳- شاخص‌های تشابه
۵۵ ۷-۲-۳- استاندارد کردن داده‌ها
۵۵ ۱-۷-۲-۳- استاندارد کردن با متغیرها
۵۶ ۲-۷-۲-۳- استاندارد کردن با مشاهدات
۵۶ ۸-۲-۳- فرضیات آنالیز خوشه‌بندی
۵۷ ۱-۸-۲-۳- نمایندگی نمونه
۵۷ ۲-۸-۲-۳- تاثیر Multicollinearity
۵۸ ۹-۲-۳- الگوریتم‌های گروه‌بندی
۵۸ ۱-۹-۲-۳- رویه‌های خوشه‌بندی سلسله مراتبی

۵۹ رویه‌های خوشه‌بندی‌غیر سلسله‌مراتبی ۲-۹-۲-۳
۵۹ مزایا و معایب روش‌های سلسله‌مراتبی ۳-۹-۲-۳
۶۰ آنالیز اجزای اصلی ۳-۳-۳
۶۰ مقدمه ۱-۳-۳
۶۰ مفهوم آنالیز اجزاء اصلی ۲-۳-۳
۶۱ کاربردهای بالقوه کاهش ابعاد ۳-۳-۳

فصل چهارم: ارزیابی تکنیک‌های کاهش ابعاد

۶۳ مقدمه ۱-۴
۶۳ ادبیات موضوع ۲-۴
۶۶ روش‌شناسی ۳-۴
۶۷ تکنیک Random Projection ۱-۳-۴
۶۷ Random Projection الگوریتم روش ۱-۱-۳-۴
۶۹ تکنیک آنالیز اجزاء اصلی ۲-۳-۴
۷۰ الگوریتم آنالیز اجزاء اصلی ۱-۲-۳-۴
۷۰ تکنیک Non-Negative Matrix Factorization ۳-۳-۴
۷۱ ارزیابی روش‌های کاهش ابعاد ۴-۴
۷۱ مجموعه داده‌های مجازی ۱-۴-۴
۷۱ مجموعه داده نرمال با ساختار خطی ۱-۱-۴-۴
۷۲ مجموعه داده حلقه‌ای با ساختار غیرخطی ۲-۱-۴-۴
۷۳ مجموعه داده متنی ۳-۱-۴-۴
۷۳ پیاده‌سازی الگوریتم‌ها بر روی مجموعه داده‌ها ۲-۴-۴
۷۸ داده‌ها با دیمانسیون بالا ۵-۴

۷۸ نتیجه گیری ۶-۴
	فصل پنجم: پیاده‌سازی متدولوژی
۸۲ ۱-۵- متدولوژی پیشنهادی
۸۳ ۲-۵- عوامل و معیارهای موثر بر بیماری دیابت
۸۴ ۳-۵- پیاده‌سازی آنالیز فاکتور
۸۶ ۱-۳-۵- تحلیل فاکتور
۸۷ ۲-۳-۵- انجام تست‌های آماری
۹۴ ۳-۳-۵- برچسب گذاری
۹۵ ۴-۵- پیاده‌سازی آنالیز اجزاء اصلی
۹۹ ۵-۵- پیاده‌سازی آنالیز خوشه‌بندی
۹۹ ۱-۵-۵- مقدمه
۱۰۱ ۲-۵-۵- اهداف آنالیز خوشه‌بندی
۱۰۱ ۳-۵-۵- طراحی روش تحقیق
۱۰۱ ۱-۳-۵-۵- شناسایی Outlier ها
۱۰۳ ۲-۳-۵-۵- تعیین معیار شباهت
۱۰۳ ۳-۳-۵-۵- استانداردسازی متغیرها
۱۰۳ ۴-۵-۵- فرضیات آنالیز خوشه‌بندی
۱۰۴ ۱-۴-۵-۵- Representativeness
۱۰۴ ۲-۴-۵-۵- Multicollinearity
۱۰۵ ۵-۵-۵- تعیین کلاسترها و پیش بینی شایستگی کلی
۱۰۵ ۱-۵-۵-۵- روش آنالیز خوشه‌بندی سلسله مراتبی

۱۰۸ شناسایی Outlier ها در تعیین خوشه‌ها
۱۱۰ ویژگی‌های هر خوشه
۱۱۱ اعتبارسنجی خوشه‌ها

فصل ششم: معرفی زمینه تحقیقات و مطالعات آتی

۱۱۵ ۱-۶- زمینه تحقیقات آتی
۱۱۵ ۲-۶- تئوری Rough Set
۱۱۵ ۳-۶- تغییر در نحوه نمونه‌گیری

۱۱۷ فهرست مراجع و منابع

۱۲۵	پیوست ۱
۱۳۳	پیوست ۲

فهرست اشکال

صفحه

۲۳	شکل ۱-۱	اثر دو نوع انسولین در بازه زمان
۲۴	شکل ۲-۱	فرآیند تبدیل داده به دانش
۳۰	شکل ۱-۲	پاسخ بدن سالم به قند خون
۳۱	شکل ۲-۲	موقعیت لوزالمعده در بدن
۳۲	شکل ۳-۲	افزایش قند خون، استون و اسیدی شدن خون
۳۳	شکل ۴-۲	پاشخ بدن افراد دیابتی نوع ۱ به گلوکز
۳۴	شکل ۵-۲	مقایسه بین الوزالمعده سالم و غیرسالم
۴۱	شکل ۶-۲	مقایسه بین انواع رژیم‌های انسولینی
۴۲	شکل ۷-۲	بررسی شماتیک هیپرگلیسمی و هیپوگلیسمی
۶۸	شکل ۱-۴	بررسی فاصله‌ها در تکرارهای مختلف در فضای کاهش ابعاد
۶۸	شکل ۲-۴	نرخ موفقیت در تصویر سازی
۶۹	شکل ۳-۴	ارتباط زمان پردازش و ابعاد پایگاه داده‌ها
۷۲	شکل ۴-۴	مجموعه داده با ساختار خطی
۷۲	شکل ۵-۴	سه مجموعه داده متحدالمركز با ساختار غیرخطی
۷۳	شکل ۶-۴	مجموعه داده با ساختار گسسته متنی
۷۵	شکل ۷-۴	نتایج حاصل از بکارگیری از سه الگوریتم بر روی مجموعه داده با ساختار خطی
۷۷	شکل ۸-۴	نتایج حاصل از بکارگیری از سه الگوریتم بر روی مجموعه داده غیرخطی
۷۹	شکل ۹-۴	نتایج حاصل از بکارگیری از سه الگوریتم بر روی مجموعه داده با ساختار متنی

- شکل ۵-۱ استفاده از نمودار Box Plot برای شناسایی داده‌های پرت ۸۵
- شکل ۵-۲ نمودار Scree Plot برای تعیین تعداد فاکتور استخراجی ۸۹
- شکل ۵-۳ نمودار سنگ ریزه ۹۸
- شکل ۵-۴ نحوه عملکرد نمونه‌های اخذ شده ۱۰۲

۳۵	جدول ۱-۲	مقایسه دیابت نوع ۱ و نوع ۲
۴۰	جدول ۲-۲	تاثیر انسولین و دامنه آن
۴۶	جدول ۱-۳	بررسی مقالات مربوط به آنالیز فاکتور در علوم مختلف
۷۶	جدول ۱-۴	مقایسه بین الگوریتم‌های PCA، RP و NMF
۷۶	جدول ۲-۴	نتایج حاصل از نمایش مجدد داده‌ها توسط سه الگوریتم
۸۷	جدول ۱-۵	نتایج تست‌های KMO & Alfa & Bartlett
۸۸	جدول ۲-۵	آمارهای توصیفی
۹۰	جدول ۳-۵	واریانس کل توضیح داده شده
۹۱	جدول ۴-۵	نتایج آنالیز فاکتور قبل از انجام چرخش
۹۳	جدول ۵-۵	نتایج آنالیز فاکتور پس از انجام چرخش
۹۴	جدول ۶-۵	برچسب گذاری و نتایج آن
۹۶	جدول ۷-۵	معیارها و عواملی که در مسئله لحاظ نشده است
۹۷	جدول ۸-۵	نتایج تست‌های KMO & Alfa & Bartlett
۹۷	جدول ۹-۵	نتایج حاصل از پیاده‌سازی آنالیز اجزای اصلی
۹۸	جدول ۱۰-۵	واریانس کلی تشریح شده
۹۹	جدول ۱۱-۵	معرفی علامت‌های بکار گرفته در فرمول تعیین نمونه
۱۰۳	جدول ۱۲-۵	بررسی نتایج تعداد داده‌های گم شده در بین ۹ متغیر X_1 تا X_9
۱۰۴	جدول ۱۳-۵	نتایج تست Multicollinearity

- جدول ۵-۱۴ مراحل طی شده جهت تعیین کلاسترها به روش سلسله مراتبی ۱۰۶
- جدول ۵-۱۵ اطلاعات لازم جهت تعیین تعداد کلاسترها ۱۰۷
- جدول ۵-۱۶ نتایج تست معنی دار بودن اختلاف بین مراکز در حالت ۸ کلاستر ۱۰۹
- جدول ۵-۱۷ نتایج تست معنی دار بودن اختلاف بین مراکز در حالت ۹ کلاستر ۱۰۹
- جدول ۵-۱۸ نتایج تست معنی دار بودن اختلاف بین مراکز در حالت ۱۰ کلاستر ۱۰۹

فصل اول

کلیات طرح تحقیق

پیشرفت‌های اخیر علوم مختلف در زمینه جمع‌آوری، ذخیره و نگهداری داده‌ها منجر به افزایش بسیار زیادی در حجم و ابعاد پایگاه داده‌ها شده است. همیشه در این نوع پایگاه داده‌ها، اطلاعاتی جهت استخراج و استفاده وجود دارد. از طرفی اطلاعات نهفته شده در این پایگاه‌های داده بسیار مفید و کاربردی خواهد بود. این اطلاعات در زمینه‌های مختلفی می‌تواند مورد استفاده قرار گیرد و مفید باشد. به عنوان مثال در تجارت جهانی، داده‌ها مبنای به وجود آمدن بازاریابی، سرمایه‌گذاری، تولید، ساخت و معاملات هستند و از این داده‌ها در جهت بررسی و تحلیل اهداف استفاده می‌شود. تحقیقات علمی یک زمینه دیگر است که می‌توان در استفاده از پایگاه داده و کسب اطلاعات از آن بهره جست.

صنایع پزشکی، جزء پنج صنعتی می‌باشد که با حجم زیاد داده و اطلاعات درگیر هستند. اطلاعات، داده‌ها و دانش ذخیره شده در این صنایع هر روز بسیار افزایش می‌یابد. در یک بررسی صورت گرفته مشخص شده است که در یک بیمارستان هر سال ۵ تریلیون بایت داده تولید و ذخیره می‌شود [۱۴]. توانائی در استفاده و استخراج دانش از این داده‌ها بسیار مهم و تعیین کننده می‌باشد. زیرا می‌تواند این صنایع را در جهت جلوگیری از بروز خطا و پشتیبانی تصمیم‌گیری‌ها یاری رساند. این امر لزوم استفاده از روش‌های پیچیده و توانمند را به عنوان یک ابزار قوی در جهت استفاده در پایگاه اطلاعات پزشکی بیان می‌کند.

داده کاوی یک رشته علمی جدید در زمینه بازاریابی و استخراج اطلاعات است که اقدام به استخراج دانش و ویژگی‌های مفید و جالب از مجموعه پایگاه داده‌ها می‌کند. به عبارت دیگر داده کاوی به عنوان یک قدم از مراحل استخراج دانش از پایگاه داده‌ها مورد استفاده قرار می‌گیرد و به رشته‌های آمار، Machine Learning، مدیریت پایگاه داده‌ها و الگو شناسی متصل شده است [۱۲].

۱-۲- ضرورت تحقیق

بر اساس اعلام سازمان بهداشت جهانی در حال حاضر دیابت سومین علت مرگ و میر در دنیاست. بر اساس آمارهای جهانی در هر ۲۰ ثانیه یک نفر به علت عدم کنترل مطلوب قند خون جان خود را از دست می‌دهد و همچنین در هر ۱۰ ثانیه یک نفر پای خود را به دلیل عوارض ناشی از دیابت از دست می‌دهد. بر اساس تحقیقات گسترده که زیر نظر و حمایت‌های مالی سازمان ملل و به جهت کمک به بیماران دیابتی صورت گرفته است، مشخص شده است که با افزایش بی-حرکی و زندگی ماشینی در جوامع جهان اول و در حال توسعه، احتمال مبتلا شدن افراد به انواع

دیابت و در نتیجه تحمیل بارهای مالی سنگین بر دولت‌ها افزایش می‌یابد. همچنین به دلیل عدم کنترل مطلوب سطح قند در حد مجاز و به علت فشار ناشی از افزایش سطح گلوکز در جریان خون صدمات ناشی از بیماری دیابت روز به روز در حال افزایش است. از این صدمات به طور خلاصه میتوان به افزایش حملات قلبی، اختلال در فعالیت کلیه، از دست دادن بینائی، قطع عضو و یا ضایع شدن اندام‌ها اشاره کرد. یکی از عواملی که در افزایش مرگ و میر ناشی از این بیماری بسیار اثرگذار بوده است بروز تدریجی عوارض این بیماری می‌باشد. عوارض دیابت می‌توانند به آرامی و در طی گذشت سال‌های متمادی به بدن آسیب برسانند، بدون آن که علامتی از خود بروز دهند. این عوارض اکثراً از بالا بودن مداوم سطح قند خون ناشی می‌شود. این عوارض به آرامی و در طی گذشت سال‌های متمادی بر گردش خون و اعصاب آسیب می‌رساند.

۱-۳- بیان مسئله

دیابت (برای درک بهتر اصطلاحات مربوط به دیابت فصل ۲ را ملاحظه فرمائید) حالتی است که در آن بدن نمی‌تواند سطح گلوکز (قند) خون را کنترل کند زیرا لوزالمعده یا نمی‌تواند انسولین ترشح کند یا سلول‌های بدن به کار انسولین مقاوم هستند. بنابراین، وقتی سطح قند خون بالا می‌رود قند نمی‌تواند وارد سلول‌های بدن شود و آنها از منبع انرژی معمولشان بی بهره می‌شوند. پاسخ بدن این است که سعی می‌کند قند اضافی را از خون خارج کند و همچنین از چربی و پروتئین (از ماهیچه) به عنوان منبع جایگزین انرژی استفاده کند. این کار باعث گسیختگی فرآیندهای بدن می‌شود و می‌تواند موجب بروز نشانه‌های دیابت شود. در بدن یک فرد سالم وقتی غذاهای کربوهیدرات‌داری مانند نان، برنج، سیب زمینی، حبوبات، یا قند خورده می‌شود، در جریان هضم غذا به قند تجزیه می‌شوند و قند از روده‌ها وارد جریان خون می‌شود و سپس وارد سلول‌های بدن می‌گردد و در آن جا به صورت سوخت می‌سوزد. یعنی به تمام بدن، از ماهیچه گرفته تا مغز نیرو می‌بخشد و منبع اصلی انرژی بدن است.

قند در کبد و در ماهیچه‌ها به شکل گلیکوژن نیز ذخیره می‌شود. دو هورمون اصلی که قند خون را کنترل می‌کنند، انسولین و گلیکوژن هستند. هر دو هورمون در لوزالمعده تولید می‌شوند. لوزالمعده غده‌ای است که پشت معده قرار دارد. در لوزالمعده سلول‌های خوشه‌ای هستند که جزایر لانگرهانس نام دارند، که در داخل آنها دو نوع سلول هورمون ساز، سلول‌های آلفا و بتا، وجود دارند. سلول‌های آلفا گلوکاگون تولید می‌کنند. در بدن انسانها، زمینه‌ی ثابتی از انسولین وجود دارد، ولی وقتی سطح قند خون بالا می‌رود، مثلاً، بعد از خوردن، انسولین زیادی به وسیله سلول‌های بتا آزاد می‌شود. انسولین مانند یک کلید عمل می‌کند و قفل سلول‌های بدن را باز می‌کند

تا قند بتواند وارد آنها شود. وقتی سطح قند خون پایین می‌آید، مثلاً پس از مدتی که غذا خورده نشود سلول‌های آلفا گلوکاگون بیشتری آزاد می‌کنند تا گلیکوژن را در کبد به گلوکز تبدیل کند. این قند یا همان گلوکز وارد جریان خون می‌شود و سطح قند خون دوباره بالا می‌رود. انسولین و گلوکاگون به طور دائم به این طریق با همدیگر طوری همکاری می‌کنند که سطح قند خون در یک حد ۴-۶ میلی مول (۷۰-۱۱۰ میلی گرم بر دسی لیتر) گلوکز در هر لیتر خون باقی بماند. در نتیجه، چه شخص کربوهیدرات زیاد بخورد، چه کم بخورد، بدن وی منبع ثابت انرژی لازم برای کارکرد مناسب را دارد.

اما وقتی بیماری قند (دیابت) در بدن وجود داشته باشد، سیستم بسیار ظریفی که قند خون را تنظیم می‌کند، دچار اختلال است. زیرا لوزالمعده یا انسولین تولید نمی‌کند یا به اندازه‌ی کافی تولید نمی‌کند و یا این که بدن به انسولین مقاوم است. در نتیجه قند نمی‌تواند وارد سلول‌ها شود تا بسوزد و انرژی تولید کند. بنابراین قند در خون باقی می‌ماند و این کار باعث می‌شود سطح قند خون از حد طبیعی و معمولی بالاتر رود. وقتی شخصی دارای سطح قند خون بالا باشد ممکن است نشانه‌هایی نظیر تکرر ادرار و تشنگی مفرط در وی وجود داشته باشد. این حالت‌ها به این علت ایجاد می‌شوند که بدن قند اضافی را به وسیله‌ی گذراندن آن از درون کلیه‌ها و از طریق ادرار خارج می‌کند (تشنگی نتیجه‌ی کمبود آب بدن است). همچنین، چون بدن نمی‌تواند برای تولید انرژی به طور مناسب از قند استفاده کند، انرژی مورد نیاز خود را با تجزیه‌ی ماهیچه‌ها و ذخیره‌های چربی به دست می‌آورد. این کار باعث کاهش وزن می‌شود و در صورت تجزیه سریع چربی، مواد شیمیایی سمی در خون به وجود می‌آید. متأسفانه، سطح قند خون بالا همیشه علامت ندارد و تنها آزمایش‌های قند خون این امکان را فراهم می‌سازد که وجود سطح قند خون بالا تشخیص داده شوند. داشتن یک سطح قند خون بالا به طور مداوم در طی چندین سال می‌تواند به نسوج بدن آسیب برساند و باعث ایجاد عوارضی در چشم‌ها، کلیه‌ها، پاها، قلب، عروق خونی و اعصاب شود. باید اشاره کرد که تا قبل از سال ۱۹۲۲ بیماران دیابتی وابسته به انسولین به علت بالا بودن میزان قند خون شانس برای زندگی دراز مدت نداشتند. اما پس از سال ۱۹۲۲ و با کشف انسولین توسط اف بانتینگ زندگی بیماران دیابتی وارد مرحله جدیدی شد. پس از آن بیماران دیابتی قبل از وعده‌های غذایی اقدام به تزریق انسولین به بدن خود می‌کنند. در حال حاضر در کشور ما بیماران دیابتی بر اساس سه فاکتور غذایی مصرفی، فعالیت بعدی و میزان قند خون در لحظه قبل از تزریق اقدام به تزریق دو نوع انسولین کوتاه اثر و اثر متوسط می‌کنند. روش تزریق در کشور بر اساس روش مدرن یا تزریقات مکرر (۲ تا ۵ نوبت در روز) است. در تعیین میزان واحدهای انسولین تزریقی عوامل و پارامترهای مهمی وجود دارد که در روش تجربی تنها بر اساس سه متغیر بالا میزان انسولین تزریقی